SPECTROSCOPY TECHNIQUES FOR HUMAN DISEASE DIAGNOSIS

by

Maria Navas-Moreno

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Physics

Department of Physics and Astronomy

The University of Utah

December 2011

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Maria Navas Moreno**

has been approved by the following supervisory committee members:

Z. Valy Vardeny	, Chair	**08/12/2011** Date Approved
Josef Prchal	, Member	**08/12/2011** Date Approved
Ajay Nahata	, Member	**08/12/2011** Date Approved
Mikhail Raikh	, Member	**08/12/2011** Date Approved
Brian Saam	, Member	**08/12/2011** Date Approved

and by **David Kieda** , Chair of

the Department of **Physics and Astronomy**

and by Charles A. Wight, Dean of The Graduate School.

ABSTRACT

Modern medicine would benefit from the pursuit of new, more specific and easier to implement diagnosis tools. In recent years, Raman scattering, surface-enhanced Raman scattering and fluorescence spectroscopy have proven to be successful diagnostic techniques for a wide range of diseases including atherosclerosis, kidney stones, bone diseases, diabetes, and a wide collection of neoplasms. Optical spectroscopy has several advantages over more traditional diagnostic methods (i.e., histopathology, quantitative PCR, etc.) such as faster data analysis, nonspecific sample preparation, nonspecific labels/reagents/antibodies usage requirements, and immediate on-site implementation.

In the present work, label-free *in vitro* fluorescence and surface enhanced Raman scattering (SERS) spectroscopy have been used to differentiate between blood cells of patients affected with myeloproliferative neoplasms (MPN) and those of healthy subjects. The SERS technique has also been applied to hemoglobin variants as well as to serum obtained from patients affected with chronic heart failure who positively or negatively responded to the seasonal influenza vaccine.

We found that spectral ratios of the background fluorescence intensity that accompanies the SERS spectra of granulocytes serve as excellent markers for the presence of MPNs. In addition, we also found expression dysregulation of two hypoxia induced factor regulated genes, which correlates with our results obtained by SERS spectroscopy assay in MPN patients and supports the presence of the Warburg effect in MPNs. We hypothesize that SERS measures metabolic change in granulocytes through two possible mechanisms: (i) Changes in dielectric properties of the environment surrounding the silver-cell interface; and (ii) changes in flavin adenine dinucleotide concentrations, which in turn changes the relative contribution of the autofluorescence to the emission spectrum. These hypotheses are supported by

SERS measurement of 2-deoxy-D-glucose incubated granulocytes, where the emission spectra show a similar behavior as observed in the SERS spectra of controls and patients.

Using SERS spectroscopy in combination with multivariate analysis (e.g., principal component analysis) and classification algorithms (e.g., support vector machines), we are able to distinguish among hemoglobin variants S, C and E and traits FS and FE, as well as seasonal influenza vaccine responders and nonresponders within a population of patients suffering from chronic heart failure.

The results presented here may have an extraordinary impact on the diagnostics of patients having Myeloproliferative Neoplasms, as well as on the health care management of patients affected with hemoglobinopathies and chronic heart failure.

A mis papás, Martha y Benito, que me han dado amor y apoyo incondicional todos los días de mi vida

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AFM	Atomic force microscopy
ATP	Adenosine triphosphate
AUC	Area under the curve
CCD	Charge couple device
cDNA	Complementary DNA
CHF	Chronic heart failure
CML	Chronic Myelogenous Leukemia
DNA	Deoxyribonucleic acid
EEC	Endogenous erythroid colonies
ELISA	Enzyme-linked immunosorbent assay
EPK	Extracellular signal-regulated kinase
EpO	Erythropoietin
EPOR	Erythropoietin receptor
ET	Essential thrombocythemia
FAD	Flavin adenosine
$FADH_2$	Reduced Flavin adenosine
H&E	Hematoxylin and eosin stain
Hb	Hemoglobin
HIF-1	Hypoxia inducible factor 1
Ig	Immunoglobulin or antibody
IRB	Institutional Review Board
JAK2	Janus kinase 2
JAK2 V617F	Janus kinase 2 gene valine to phenylalanine substitution at residue 617
LDH	Lactate dehydrogenase
LOH	Loss of heterozygosity
LSPR	Localize surface plasmon resonance
MAPK	Mitogen-activated protein kinase
MPN	Myeloproliferative neoplasm
MDP	Myeloproliferative Disorder
MPS	Myeloproliferative Syndrome
MVA	Multivariate analysis

NA	Numerical Aperture
NAD	Nicotinamide adenine dinucleotide
NADH	Reduced Nicotinamide adenine dinucleotide
PC	Principal component
PCA	Principal component analysis
PCR	Polymerase Chain Reaction
PMF	Primary myelofibrosis
PDK1	Pyruvate dehydrogenase kinase, isozyme 1
PMT	Photomultiplier tube
PV	Polycythemia vera
qRT-PCR	Reverse Transcription Real Time Polymerase Chain Reaction
RCM	Red cell Mass
RMS	Root mean square
RNA	Ribonucleic Acid
ROC	Receiver operating characteristic
RRS	Resonance Raman scatering
SEM	Scanning electron microscopy
SERS	Surface enhanced raman scattering
SGH	Second Harmonic Generation
SLC2A1	Solute carrier family 2 member 1 (or Glucose transporter 1)
SNOM	Scanning near-field optical microscopy
SPP	Surface plasmon polariton
ssDNA	Single strain DNA
SVM	Support vector machine
TCA	Tricarboxylic Acid
TSG	Tumor suppressor gene
VEGF-A	Vascular endothelial growth factor A
WHO	World health Organization

ACKNOWLEDGMENTS

Doing this PhD is, by no doubt one of the toughest things I have done. Doing a PhD can drain you both mentally and physically. That is why without the people around me I couldn't have done it.

I would like to start thanking both Professor Valy Vardeny and Doctor Josef Prchal for believing in me and for taking a risk with this project. Valy and Joe, I really value your mentoring and the time you have given to me and this project. What I have learned from you is invaluable, and its something that will be with me for the rest of my life.

To Professor Sabina Swierczek, thank you for all the help, time and special interest in this research project. I am very grateful of your support. To Kimberly Hickman, thank you for your invaluable help with sample acquisition and IRB protocols. In the Prchal's Lab, Susan and Irma, thank you for your time and patience. I hope I didn't cause much trouble in the lab. In Vardeny's group, Ella, Bhoj, Bill, Tek, Tho and Sheng, thank you for patiently listening to my talks in group meetings. Thanks to Doctor Nick Borys in Lupton's group for help with measurements, very useful scientific discussions and proof reading this document.

To all the staff in the physics department: Heidi, Jolene, Kathy, Vicky, Ed, Kathrine, Wayne, Matt, Harold, Brad and Mary, thank you! You made my life so much easier. Randy and Leonard, thank you so much for all the knowledge, time, patience and technical support. I would like to specially thank Jackie. You are my beloved friend and since the first day I got to Utah you have been there supporting me.

I would like to thank Professor Orly Vardeny in the School of Pharmacy at the University of Wisconsin-Madison for her time and useful discussions.

To all my friends in Salt lake, thank you for making me forget, for longer than I

should have, about physics and the pain of doing a PhD, for giving me the chance to make cakes to celebrate birthdays and for being there all along the way. I have to make a pause here and thank three very special people: Su, Golda and Jose. Your love, advice, good and bad jokes, lunch and coffee time made this journey so much more pleasant, I am going to miss you with all of my heart.

To my grandma, aunts, my cousins, my siblings, my nephews and my niece, Gabriel and my friends in Colombia and abroad. Thank you for putting up with me being far away and still loving me.

To Andrés: Thank you for giving me hope back and for reminding me every day to believe in my self. Without that last push I don't think I could have made it. Thank you for listening to me, even when I don't make sense. Te amo.

Last but not least I would like to thank my parents, friends and confidants, for giving unconditional love and support and for making me feel like everything is possible. No hay nada en el mundo que exprese lo mucho que los quiero y lo que significan para mi.

CHAPTER 1

INTERACTION OF LIGHT WITH MATTER

Assays that offer better detection of a disease or a dysfunction are highly desired. Such assays should provide prompt, quantitative and accurate assessments of the disease while ideally being inexpensive, noninvasive, reagent free and easily and quickly implementable. Optical spectroscopy techniques have fulfilled some of these requirements and in contrast with current diagnostic methods, they are capable of simultaneously linking morphological and molecular properties of tissues or cells.

Optical spectroscopy, especially autofluorescence (i.e., fluorescence of intrinsic fluorophores) and Raman scattering, have been extensively studied for medical diagnosis in the last 10 years, mainly due to the development of relatively inexpensive and very efficient light detection devices and the advancements in statistical and computational methods that allow processing of the increasing amount and complexity of the data collected.

In the present chapter some of the underlying mechanisms governing the interaction of light with matter are introduced. Emphasis is put on fluorescence, Raman and surface enhanced Raman scattering (SERS) since these methods constitute the main optical spectroscopic techniques used in the present work. Chapter 2 provides a brief introduction to the molecular pathogenesis of the diseases and dysfunctions studied and general aspects of the diagnostic techniques currently used. The experimental and statistical tools used for the research presented here are described in Chapter 3 followed by the results myeloproliferative neoplasms (Chapter 4) and, hemoglobin variants and blood serum (Chapter 5). A brief summary and future directions are presented Chapter 6.

1.1 Luminescence

The emission of light from excited electron states of any substance is dubbed luminescence. Depending on the nature of the excited state, luminescence is divided into two categories: fluorescence and phosphorescence. They differ from each other in that fluorescence emission occurs from a singlet excited state, whereas phosphorescence emission occurs from a triplet excited state.

Luminescence is preceded by the absorption of photons by atoms or molecules, and it is only one way in which the excited electrons can relax to the ground state. An atom or a molecule can absorb a photon with energy that is equal to the energy necessary to take the atom or the molecule to the next excited state (ΔE).

Due to line-width broadening, ΔE is not a single value but a distribution of energy values whose shape depends upon the following factors: 1) The uncertainty principle, otherwise known as **natural broadening**, which depends on the excited state lifetime; 2) **Doppler broadening** where the velocity distribution of the atoms or molecules results in a distribution of absorption frequencies that follows the Maxwell's statistical distribution function; 3) **Collision broadening** in which the atomic or molecular energy levels are shifted due to interparticle collisions; and 4) **power broadening**, which takes place above a certain excitation intensity threshold where nonlinear optical phenomena start to occur.

In addition to the line-width broadening, transitions from the ground state to the excited state can take place from one of the vibrational sublevels in the ground state to one of those associated with the excited state. As a result photons with a broad spectrum of energies can be absorbed. Similarly, the luminescence spectrum can also be symmetrically broadened.

There are many competing processes by which the excited atom or molecule can return to the ground state. Fluorescence competes with phosphorescence as well as with other relaxation mechanisms of relaxation that do not involve the emission of photons, namely nonradiative transitions:

- **Internal conversion** by which the atom or the molecule is able to relax by converting the electronic energy into vibrational and rotational energy.

- **Energy transfer** by which the atom or the molecule transfers the excess energy to another by either collision or resonance transfer energy (RET)

- **Intersystem crossing** by which the excited electron's spin flips and the molecule goes from the singlet to the triplet manifold, thus further internal conversion or the emission of a photon (phosphorescence) may proceed.

The fate of an excited state is summarized in a Jablonski-type diagram as schematically shown in Figure 1.1.

The intensity at which the substance will fluoresce is determined by the competition among the processes describe above, and consequently the quantum yield can be defined as the ratio of emitted photons to absorbed photons, given by the following equation:

$$\Phi = \frac{K_f}{\sum_i K_i} \tag{1.1}$$

where K_f us the dissipation rate of fluorescence and K_i are the dissipation rates of various processes.

1.2 Raman scattering

Scattering is one of the phenomena by which light and matter interact. Elastic scattering (i.e., Rayleigh and Mie scattering) is the case when the scattered photons have the same energy as the incoming ones. In contrast, Raman scattering is inelastic light scattering in which the impinging photons lose (or gain) energy to (or from) the scattering medium. Unlike absorption, scattering does not require the energy of the incident photon, E_{in}, to be equal to the energy difference between electronic energy levels ΔE. However when E_{in} becomes closer to ΔE, the Raman cross-section is enhanced, a phenomenon known as resonant Raman scattering (RRS).

Stokes raman scattering refers to the case in which the photon loses energy to the scattering material, whereas **anti-Stokes Raman scattering** is the case in which the photon gains energy from the scattering medium. The amount of energy

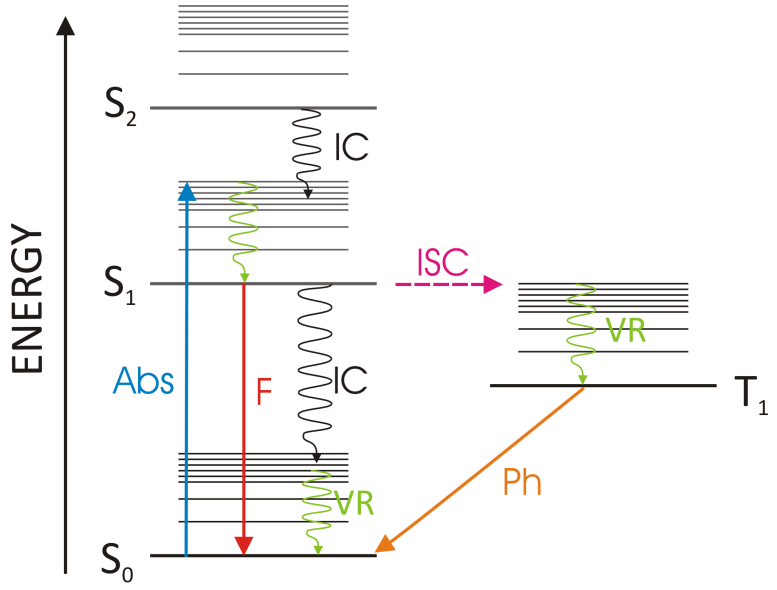


Figure 1.1: Jablonski diagram of the possible transitions following photon absorption of a photon. S=Singlet state, T=Triplet State, Abs=Absorption, F= Fluorescence, Ph=Phosphorescence, VR= Vibrational Relaxation, IC= Internal Conversion and ISC= Intersystem crossing to the triplet manifold.

lost (or gained) is equal to the the energy difference between two vibrational states of the molecules in the medium.

The origin of the scattered light is the oscillating electric and magnetic moments induced in a molecule by the electromagnetic field of the incident light. In the classical theory framework, the induced electric dipole of a molecule in the presence of an external electric field can be written as a converging series:

$$\vec{p} = \vec{p}^{(1)} + \vec{p}^{(2)} + \vec{p}^{(3)} + \cdots \tag{1.2}$$

The relation between the first three induced dipole moments and $\vec{E}$, the field of the incident monochromatic plane wave, is given by:

$$\vec{p}^{(1)} = \bar{\bar{\alpha}}\vec{E} \tag{1.3}$$

$$\vec{p}^{(2)} = \frac{1}{2}\bar{\bar{\beta}}\vec{E}\vec{E} \tag{1.4}$$

$$\vec{p}^{(2)} = \frac{1}{6}\bar{\bar{\gamma}}\vec{E}\vec{E}\vec{E} \tag{1.5}$$

where $\bar{\bar{\alpha}}$, $\bar{\bar{\beta}}$ and $\bar{\bar{\gamma}}$ are the second rank (2×2) polarizability tensor, third rank (3×3) hyperpolarizability tensor and the fourth rank (4×4) second hyperpolarizability tensor, respectively.

The Raman scattering frequency is obtained from the first induced dipole moment $\vec{p}^{(1)}$. The polarizability tensor $\bar{\bar{\alpha}}$ is a function of the nuclear coordinates. Nuclei vibrate about their equilibrium positions, and such vibrations can be expressed in terms of normal coordinates $\vec{q}$, and each component of the polarizability tensor α_{ij} can be expressed as an expansion around the equilibrium position [1]:

$$\alpha_{ij} = (\alpha_{ij})_0 + \sum_n \left(\frac{\partial \alpha_{ij}}{\partial q_n}\right)_0 q_n + \frac{1}{2}\sum_{n,m}\left(\frac{\partial^2 \alpha_{ij}}{\partial q_n \partial q_m}\right)_0 q_n q_m + \cdots \tag{1.6}$$

Assuming both mechanical[1] and electrical [2] harmonicity, eq. 1.6 takes the form:

$$\alpha_{ij} = (\alpha_{ij})_0 + (\alpha'_{ij})_n q_{n0}\cos(\omega_n t + \delta_n) \tag{1.7}$$

where $q_n = q_{n0}\cos(\omega_k t + \delta_k)$ due to mechanical harmonicity and $(\alpha'_{ij})_n = \left(\frac{\partial \alpha_{ij}}{\partial q_n}\right)_0$.

The incident monochromatic plane wave frequency dependence is given by:

$$\vec{E} = \vec{E}_0\cos(\omega_1 t) \tag{1.8}$$

and the first induced dipole moment is then expressed as:

$$\vec{p}^{(1)} = \bar{\bar{\alpha}}_0\vec{E}_0\cos(\omega_1 t) + \bar{\bar{\alpha}}'_n q_{n0}\vec{E}_0\cos(\omega_n t + \delta_n)\cos(\omega_1 t) \tag{1.9}$$

---

[1]Vibrations are described by the simple harmonic oscillator.

[2]Only the first two terms of the expansion of the polarizability are considered.

Using the trigonometric identity $\cos A \cos B = \frac{1}{2}(\cos(A+B) + \cos(A-B))$, eq. 1.9, can be rewritten in the following form:

$$\vec{p}^{(1)} = \vec{p}^{(1)}(\omega_1) + \vec{p}^{(1)}(\omega_n + \omega_1) + \vec{p}^{(1)}(\omega_n - \omega_1) \tag{1.10}$$

In eq. 1.10, the first term $\vec{p}^{(1)}(\omega_1)$ correspond to Rayleigh scattering and the other two terms, $\vec{p}^{(1)}(\omega_n+\omega_1)$ and $\vec{p}^{(1)}(\omega_n-\omega_1)$ correspond to anti-Stokes and Stokes Raman scattering, respectively.

The intensities of the Rayleigh (eq. 1.11) and Stokes and anti-Stokes Raman (eq. 1.12) scattered light correspond to the intensities of electric dipoles at frequencies ω_1, $\omega - \omega_1$ and $\omega + \omega_1$, respectively, where ω is the frequency of a molecular vibration. For one dipole that is:

$$I_{Rayleigh} = \frac{\omega_1^4 \left(\bar{\bar{\alpha}}_0 \cdot \vec{E}_0(\omega_1)\right)^2 \sin^2\theta}{32\pi^2\epsilon_0 c^3} \tag{1.11}$$

$$I_{Raman} = \frac{(\omega \pm \omega_1)^4 \left(\bar{\bar{\alpha}}_{Raman} \cdot \vec{E}_0(\omega_1)\right)^2 \sin^2\theta}{32\pi^2\epsilon_0 c^3} \tag{1.12}$$

In the case of bulk matter, the intensities of the Stokes and anti-Stokes scattered light are proportional to the average number of molecules at the vibrational ground and excited states, respectively. In thermal equilibrium the vibrational population of the ground (eq. 1.13) and excited (eq. 1.14) states is given by the Boltzmann distribution.

$$\bar{n}_g = N\frac{1}{\sum_r e^{-\frac{\hbar\omega_r}{kT}}} \tag{1.13}$$

$$\bar{n}_s = N\frac{e^{-\frac{\hbar\omega_s}{kT}}}{\sum_r e^{-\frac{\hbar\omega_r}{kT}}} \tag{1.14}$$

where k is the Boltzmann constant, $\hbar$ is the reduced Planck constant, T is the temperature and N is the total number of molecules. From eqs. 1.12, 1.13 and 1.14 it is found that the ratio of anti-Stokes to Stokes Raman scattered light at the i-th vibration mode is proportional to the ratio of the forth power of the scattered frequencies and to the ratio of the excited and ground states populations:

$$\frac{I_{AS}}{I_S} \propto \frac{(\omega+\omega_1)^4}{(\omega-\omega_1)^4} e^{-\frac{\hbar\omega_s}{kT}} \tag{1.15}$$

The proportionally constant for relation 1.15 is determined by the polarizability tensors for Stokes and anti-Stokes Raman scattering, which under nonresonant conditions are the same thus making relation 1.15 an equation. Consequently, from the Raman scattering spectrum one can calculate the temperature of the sample as shown in Figure 1.2, where the temperature of C_{60} molecules was calculated for different illumination conditions from the ratio, I_{AS}/I_S.

The polarizability tensors $\bar{\bar{\alpha}}_0$ and $\bar{\bar{\alpha}}_{Raman}$ carry information about the molecule and determine whether Rayleigh or Raman scattering are observable, respectively. Since all molecules are polarizable to some extent, $\bar{\bar{\alpha}}_0$ always has nonzero components; thus Rayleigh scattering is observable from all molecules. On the other hand, from eq. 1.9 it is evident that for Raman scattering to be observable, $\bar{\bar{\alpha}}_{Raman}$ must have a nonzero gradient (evaluated at the equilibrium position). This condition imposed on $\bar{\bar{\alpha}}_{Raman}$ implies that not all vibrational modes are Raman active. Some of the components of $\bar{\bar{\alpha}}_{Raman}$ can still be zero. For those modes the derivative of the permanent dipole moment at the equilibrium position is finite and thus infrared active [1]. If the molecule is centro-symmetric, there is a complimentary relation between modes that are Raman-active and those that are IR-active.

Raman scattering spectrum is highly molecular specific, thus allowing identification of the active molecules. Also, since Raman scattering intensity is proportional to the number of molecules, relative concentrations of substances can be measured in some circumstances. This has enabled Raman scattering to become an important tool in chemical analysis, specially in the pharmaceutical industry [3].

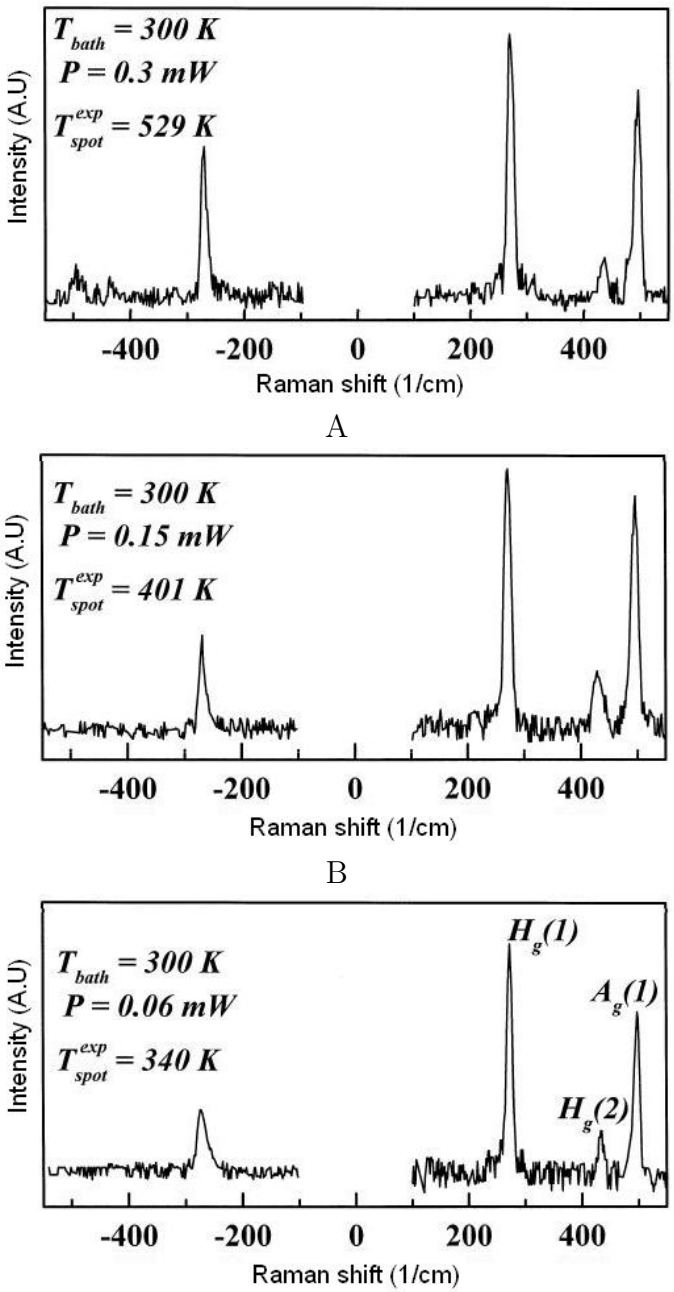


Figure 1.2: Raman spectra of solid C_{60} at room temperature and for different laser power densities. Reprinted with permission from [2]. Copyright 1998, Elsevier.

The Raman scattering process described above is dubbed spontaneous and is a linear process. Eq. 1.12 shows that spontaneous Raman scattering has an intensity that depends linearly with the excitation intensity (i.e., the square of the incoming electric field). From eq. 1.12 the power of the Stokes process from N active scatterers may be expressed as:

$$P_S(\omega_s) = N\sigma I(\omega) \tag{1.16}$$

where σ is the Raman scattering cross section and $I(\omega)$ is the intensity of the incident beam.

Typical Raman scattering cross sections are between $10^{-31}\frac{cm^2}{molecule}$ and $10^{-29}\frac{cm^2}{molecule}$ making spontaneous Raman scattering an inherently weak process, and for practical applications, enhancement is necessary. SERS is an enhancement technique where the probed molecules are close to a metallic nanoparticle. The theory of SERS is summarized in the following section.

1.3 Surface enhanced Raman spectroscopy

Most of the features observed from the enhancement of Raman scattering near a metallic nanoparticle can be explained by two effects [4–6]: (i) ***electromagnetic enhancement*** due to the excitation of localized surface plasmons in the metal structure and (ii) ***chemical enhancement*** due to adsorption of molecules onto the metallic surface.

Electromagnetic enhancement originates from the ability of subwavelength metallic particles to sustain oscillating surface plasmons induced by impinging electromagnetic waves. The surface plasmons are collective oscillations of the free, or almost free, electrons on top of the stationary ionic nuclei background. Localized surface plasmons are not to be confused with surface plasmon polaritons. To disambiguate these two concepts, a brief explanation of the former is given followed by the explanation of localized surface plasmons, which are responsible for *electromagnetic enhancement* of Raman scattering.

Surface plasmon polaritons (SPP) are electromagnetic waves propagating along the flat interface between a dielectric and a conductor. They decay exponentially in the direction perpendicular to the metal surface. Using Maxwell's equations and the continuity of the electric and magnetic fields, one can show that only transverse magnetic (TM) modes excite SPPs. For TM polarization the dispersion relation of SPPs is given by:

$$\beta = k_0 \sqrt{\frac{\epsilon_1 \epsilon_2}{\epsilon_1 + \epsilon_2}} \tag{1.17}$$

where β is the propagation constant of the SPP, k_0 is the wave vector in the dielectric, and ϵ_1 and ϵ_2 are the frequency dependent dielectric constants of the metal and dielectric, respectively. From eq. 1.17 one can see that the SPP momentum exceeds the momentum of a photon in the dielectric ($k_0 \epsilon_1^{1/2}$) for every frequency ω. Thus, in order to excite SPPs, phase matching is required.

In the case of small metallic structures, the high surface curvature leads to non-propagating excitations of the conducting electrons dubbed ***localized surface plasmons***. The surface curvature acts as an effective restoring force that allows resonances and also renders phase matching conditions unnecessary [6]. The main features of localized surface plasmon resonances (LSPR) can be understood by studying the case of a subwavelength spherical metallic particle. Assuming that a, the radius of the particle, is smaller than the wavelength λ_0 of the impinging electromagnetic wave, it is valid to use the electrostatics approach, i.e., solve for Laplace's equation of a metallic particle in an uniform electric field (see Figure 1.3). The solution yields the following electric potentials inside and outside the metallic nanoparticle:

$$\phi_{in} = -\frac{2\epsilon_m}{\epsilon + 2\epsilon_m} E_0 r \cos\theta \tag{1.18}$$

$$\phi_{out} = -E_0 r \cos\theta + \frac{\epsilon - \epsilon_m}{\epsilon + 2\epsilon_m} E_0 a^3 \frac{\cos\theta}{r^2} \tag{1.19}$$

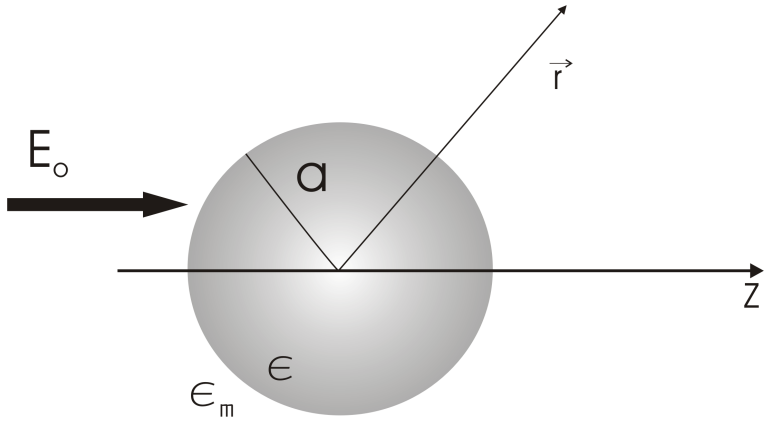


Figure 1.3: Metallic sphere placed in a uniform electric field.

Eq. 1.19 is the superposition of the applied field and the dipole located at the center of the sphere, and it can be rewritten in terms of the dipole moment $\overrightarrow{P}$ as:

$$\phi_{out} = \phi_0 + \frac{\vec{P} \cdot \vec{r}}{r^3} \tag{1.20}$$

where $\overrightarrow{P} = a^3 \frac{\epsilon - \epsilon_m}{\epsilon + 2\epsilon_m} \overrightarrow{E_0}$. Thus, the polarizability of a small metallic particle is the given by:

$$\alpha = a^3 \frac{\epsilon - \epsilon_m}{\epsilon + 2\epsilon_m} \tag{1.21}$$

From eq. 1.21 it is evident that α exhibits a resonance as $\epsilon \to -2\epsilon_m$. For a Drude metal (i.e., free electron gas) the dielectric constant has the form $\epsilon_0(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + i\gamma\omega}$ where $\omega_p = \frac{ne^2}{\epsilon_0 m}$ is the plasma frequency. Taking interband transitions into account the dielectric constant becomes $\epsilon(\omega) = \epsilon_{IT} + \epsilon_0(\omega)$ and the polarization assumes the form:

$$\alpha = \frac{a^3(\epsilon_{IT}\omega^2 - \omega_p^2) + i\omega\gamma\epsilon_{IT}}{((\epsilon_{IT} + 3)\omega^2 - \omega_p^2) + i\omega\gamma(\epsilon_{IT} + 3)} \tag{1.22}$$

Eq. 1.22 has a resonance with width $\sim \gamma(\epsilon_{IT}+3)$. If the metal is a poor conductor, γ is large and the quality of the resonance goes down. Resonance quality is also reduced for metals whose dielectric properties are highly modified by interband transitions [5]. Interband contributions are larger in gold than in silver, thus explaining the larger SERS enhancement observed in silver [5]. Additionally, in the visible range the real part of $\epsilon(\omega)$ is fairly close to $-2\epsilon_0$ for both gold and silver (see Figure 1.4), but the imaginary part is higher in gold. This implies that in the visible range the enhancement factor in silver is larger.

LSPRs lead to the enhancement of the electric field in the vicinity of a metallic nanostructure as demonstrated for the case of a metallic nanosphere. From eqs. 1.21 and 1.22 the expected dependence of LSPR on the geometry of the particle is not evident. To gain a little more insight of the LSPR one can solve for the polarizabilities of an ellipsoid along the principal axis. Eq. 1.23 shows such polarizabilities where L_i is a geometrical factor and $\sum_i L_i = 1$. Notice that for a spheroidal particle (i.e., an ellipsoid with two identical principal axis) only two spectrally resolved plasmon

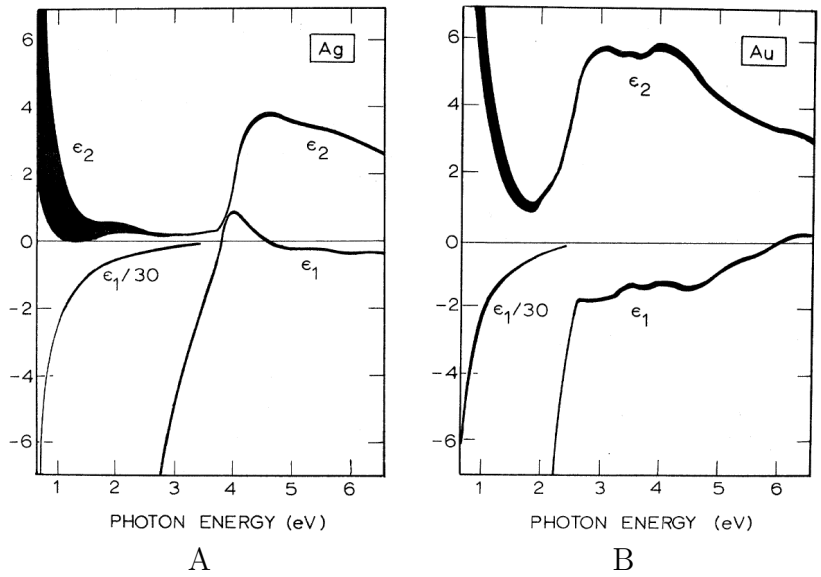


Figure 1.4: The real ϵ_1 and imaginary ϵ_2 dielectric constants of **A:** Silver and **B:** Gold normalized to ϵ_0. Reprinted with permission from [7]. Copyright 1972, American Physical Society.

resonances are observed as two plasmon resonances are degenerate.

$$\alpha_i = a_1 a_2 a_3 \frac{\epsilon(\omega) - \epsilon_m}{3\epsilon_m + 3L_i(\epsilon(\omega) - \epsilon_m)} \tag{1.23}$$

The resonance of spheroids along the major axis are red-shifted compared to the plasmon resonances of a sphere with the same volume. This is a very attractive property of LSPR as it allows tunability of the resonances by changing the aspect ratio of the particles.

In the presence of the enhanced electromagnetic field, both the incoming and the scattered light are enhanced, thus the power of the Raman scattered light can be expressed as:

$$P_S(\omega_s) = N\sigma[G(\omega_0)E_0(\omega)]^2 G^2(\omega_s) \tag{1.24}$$

where $G(\omega)$ is the enhancement factor defined as $G(\omega) = \frac{E_M}{E_0}$ and E_M is the magnitude of the enhanced field close to the metallic nanostructure.

In the case of the metallic sphere from eq. 1.19 the enhanced electric field at a distance $\vec{r}$ from the sphere is given by $\vec{E_M} = \vec{E_0} + \frac{\epsilon-\epsilon_m}{\epsilon+2\epsilon_m} * \frac{a^3}{r^3}\vec{E_0}$, thus the factor is:

$$G(\omega) = \frac{E_M}{E_0} = 1 + \frac{\epsilon - \epsilon_m}{\epsilon + 2\epsilon_m} * \frac{a^3}{r^3} \tag{1.25}$$

that near the resonant condition $\epsilon \to -2\epsilon_m$ becomes:

$$G(\omega) = \frac{\epsilon - \epsilon_m}{\epsilon + 2\epsilon_m} * \frac{a^3}{r^3} \tag{1.26}$$

Examples of the enhancement factor under resonant conditions for various geometries are displayed in Figure 1.5. Notice that for the bipiramidal geometry, the enhancement factor is the highest at the tips. This concentration of electric field is

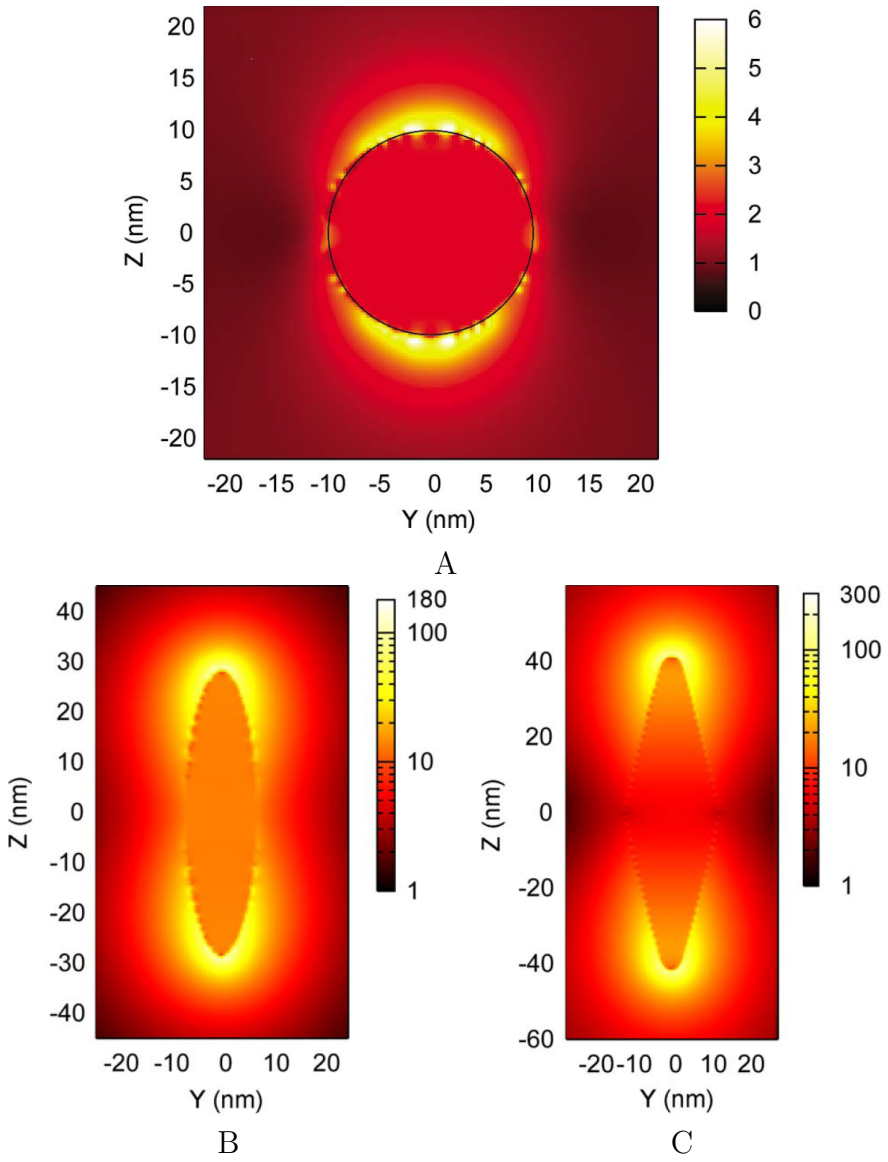


Figure 1.5: The field enhancement factor at the surface plasmon resonance of **A:** a gold sphere **B:** a gold spheroid and **C:** a gold bipiramid calculated by finite difference time domain (FDTD) method. Polarization is along the z direction. Reprinted with permission from [8]. Copyright 2007, American Physical Society.

due to the increase of induced charge density at the sharper features of the geometry, known as *lighting rod* effect.

An LSPR enhances not only Raman but other linear and nonlinear processes. Such is the case of fluorescence and second harmonic generation. In the case of fluorescence, the enhancement results as a consequence of the modification of the excitation and the emission processes alike. In the surroundings of a metallic structure the fluorescent molecule (i.e., fluorophore) can be excited by the enhanced EM field and the metallic surface can serve to direct the emission towards the detector as well, thus increasing the overall detected intensity of the emitted light. Moreover the metallic nanostructure modifies the relaxation processes of the fluorophore; close proximity to the metal surface increases the rate of nonradiative process thus quenching fluorescence. For larger distances, the rate of radiative processes is higher than nonradiative ones due to modifications of the environment, and therefore fluorescence is enhanced. The increase of the radiative rate is a consequence of the reduction of the fluorescence lifetime (Purcell effect) that increases the emission quantum yield [9]. Since the surface of any material lacks inversion symmetry, which is a necessary condition for second harmonic generation (SGH), the metallic nanostructures thus emit frequency doubled light. The enhancement of SGH is a direct consequence of the enhancement of the excitation and emitted electric field due to localization [6, 10].

An additional mechanism for the emission enhancement is the so called ***chemical enhancement***. Chemical enhancement encompasses a handfull of observations that the electromagnetic theory is incapable of explaining. Among them, molecules with almost identical Raman cross sections have SERS spectra that vary both in intensity and the amount of shift that Raman peaks display. Such effects are observed in SERS spectra of CO and N_2 [11] and ethane/methane and ethylene/benzene [4]. Additionally, EM theory predicts that the SERS spectrum should be more or less an enhanced copy of the normal Raman spectrum. In reality, however, forbidden vibrational bands are observed with SERS whereas bands expected to have high intensity are abscent. Finally there is a disagreement between the predicted enhancement factor between EM theory ($\sim 10^{11}$) and the actually observed factor (10^{14}, [12, 13]).

These observations can be explained by a resonant process due to the adsorption of the probed molecules onto the metallic nanostructure surface [5].

1.4 Properties of nanoparticles

As described above, the LSPR frequency depends not only on the metal but also on the size and shape of the particle as well as the dielectric properties of the surrounding medium [14]. Additionally, interparticle interactions also change the optical properties, and will be explained in more detail below.

The condition for the LSPR dictates that $\epsilon \to -2\epsilon_m$; thus the electric properties of the surrounding media are capable of moving the resonance towards higher or lower wavelengths. Such changes are reflected in experiments done by Evanoff *et al.*, where Ag nanoparticles were placed in five different media of different dielectric constants [15] (see Figure 1.6).

The LSPR also depends on the shape and size of the nanoparticles. An example of this is shown in Figure 1.7A, where the extinction spectra of spherical, pentagonal and triangular particles of comparable sizes are shown. Figure 1.7B shows the extinction spectra for spherical particles with different diameters. As the particles increase in

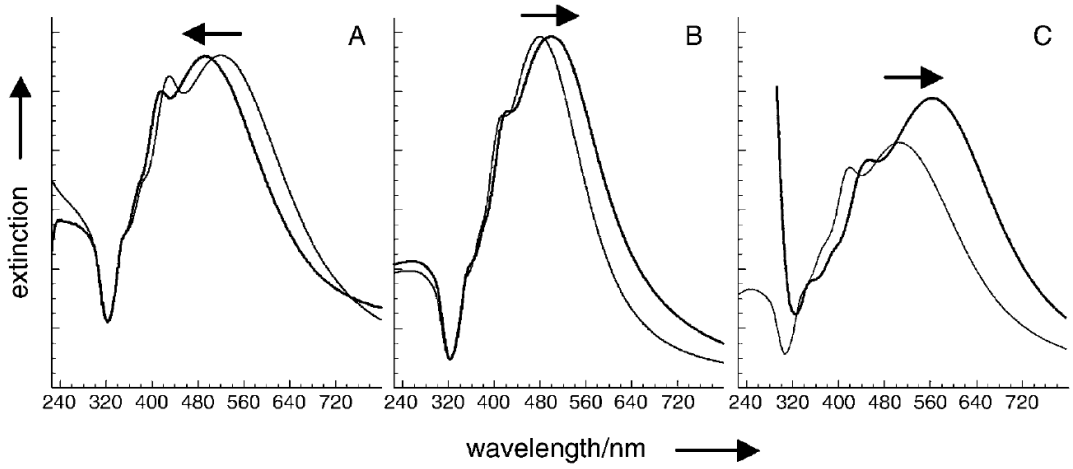
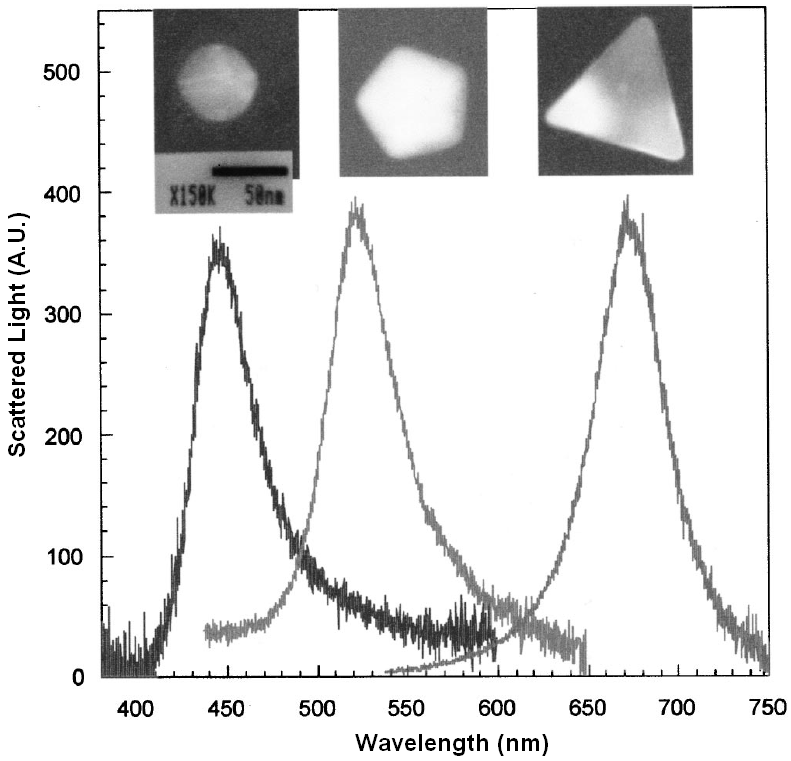
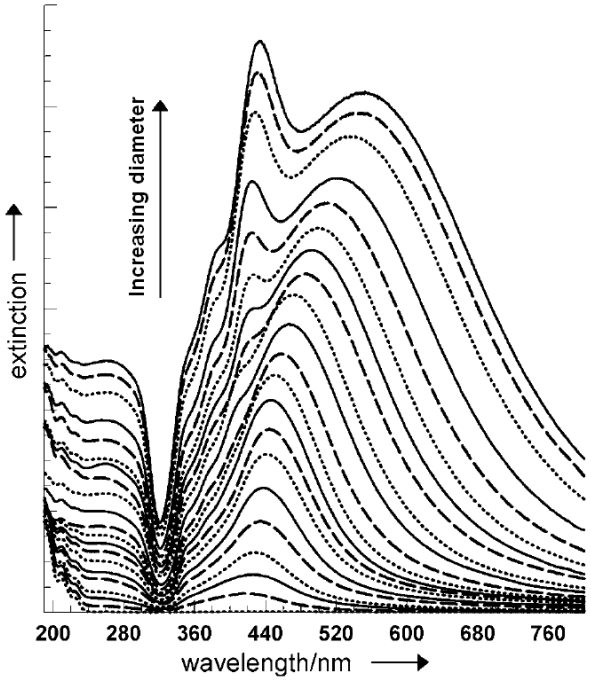


Figure 1.6: Extinction spectra of silver nanoparticles in solution before (thin line) and after (thick line) surrounding dielectric medium was changed from **A:** isopropanol to teflon, **B:** water to silica and **C:** water to titania. Reprinted with permission from reference [15]. Copyright 2007, Springer.



A



B

Figure 1.7: Extinction spectra of **A:** spherical, pentagonal and triangular silver nanoparticles particles. Reprinted with permission from [16]. Copyright 2002, American Institute of Physics; and **B:** spherical silver nanoparticles with 20 different diameters. Reprinted with permission from [15]. Copyright 2007, Springer.

diameter, the maximum in the LSPR is red-shifted. The second peak that appears at $\sim 420nm$ is attributed to the quadrupole [15].

For many applications, single nanoparticles are not used. Instead metallic aggregates, which have different optical properties, are utilized. In metallic aggregates, the individual particles couple through dipole-dipole interactions and generate plasmon modes that cover the a larger protion of the aggregate. Such modes cover a broader frequency region as shown in Figure 1.8 [14, 17, 18].

In the last 30 years or so, thanks to the development of technologies to manufacture nanoparticles in a controlled and reproducible way, many applications in medical therapeutics, diagnosis and drug delivery have flourished. Some applications, such as those aimed at cancer therapy and magnetic resonance imaging (MRI) contrast, have been already approved by the Food and Drug Administrations (FDA) and are currently in clinical use [19].

1.5 Thermal effects and photochemical processes

Following light absorption, molecules are left in an excited stated that can relax back to the ground state through radiative and nonradiative processes. The nonradiative process proceeds by dissipating the energy from the excited state as heat or by producing a chemical reaction (i.e., photochemical reactions).

The thermal effects are attributed internal conversion, intercrossing system, and/or vibrational relaxation (i.e., conversion of energy to vibrational energy). A notorious thermal effect is *carbonization* which occurs in organic compounds such as polymers or biologically relevant molecules like proteins, lipids, polysaccharides and nucleic acids. Through the heat induced cleavage of bonds and rearrangement, these materials are converted into amorphous carbon-like structures. Other thermal effects are slow compared to photochemical process, which can be three to four orders of magnitude faster.

Photochemical processes include photoaddition, photofragmentation, photooxidation, photohydration, *cis-trans* isomerization and photorearrangement [20]. Some examples [20] are the photodymerization of thymine, light induced protein-DNA crosslinking, rivoflavin photofragmentation, photooxidation of cholesterol, photoiso-

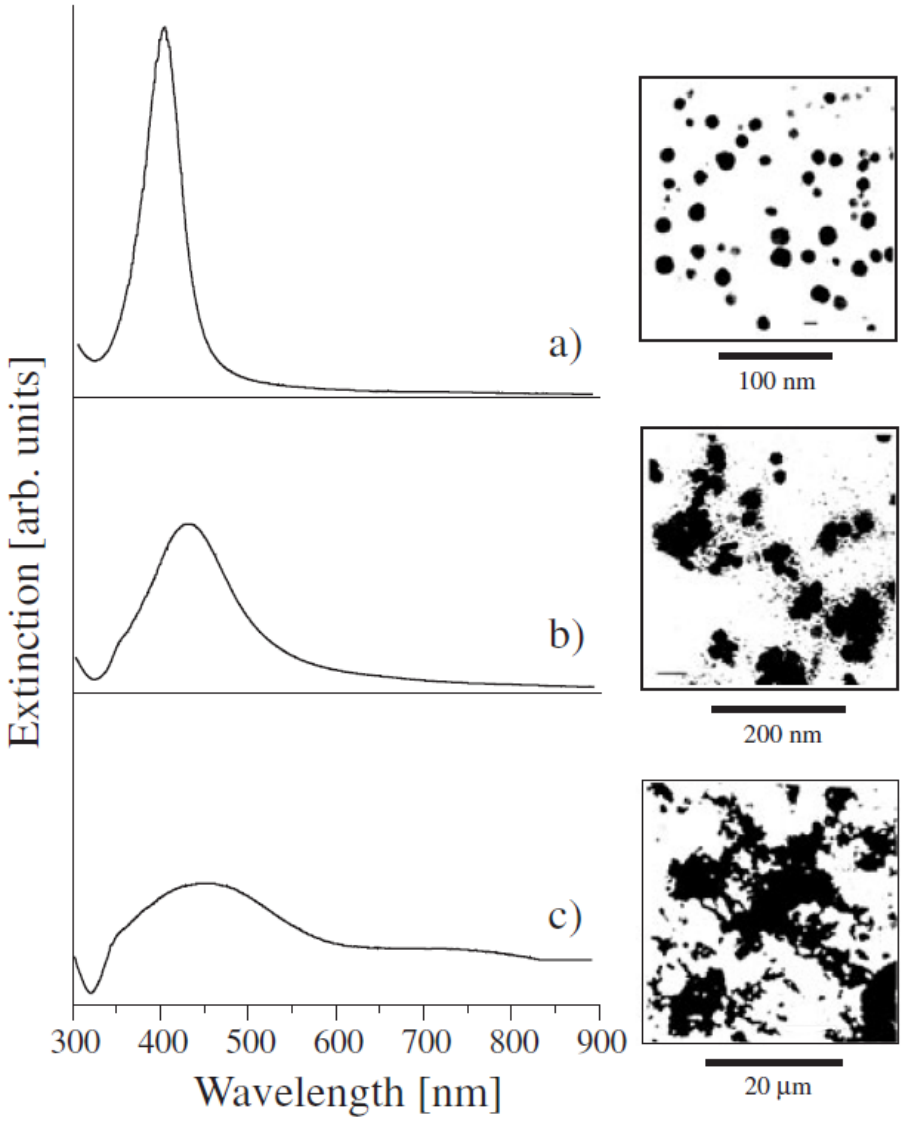


Figure 1.8: Extinction spectra for aggregates at various aggregation states as shown by the SEM images. Reprinted with permission from [17]. Copyright 2002, IOP.

merization of 11-*cis*-retinal and the rearrangement of 7-dehydrocholesterol to vitamin D upon UV radiation. See Figures 1.9 and 1.10.

1.6 Interaction of light with cells and tissue

The interaction of light with biological samples, such as cells and tissue, leads to fluorescence and light scattering processes as well as photochemical processes. Some of the molecules that make up biological specimens fluoresce. Such molecules are called endogenous or intrinsic fluorophores. The absorption and emission spectra of representative endogenous fluorophores are depicted in Figure 1.11. In the case of proteins (e.g., collagen and elastin), the fluorescence originates with the aromatic amino acids tryptophan, tyrosine and phenylalanine; the emission of tryptophan is particularly sensitive to its surroundings; hence it is used as a reporter of protein conformational changes [21]. Another important group of endogenous fluorophores are the metabolism coenzymes called Nicotinamide adenine dinucleotide-Reduced ($NADH$) and Flavin adenine dinucleotide (FAD). These coenzymes help cellular processes by carrying electrons from one reaction to another. Thus they are essential for cellular metabolism. NADH and FAD exist also in the oxidized NAD^+ and reduced form $FADH2$, respectively, which are not fluorescent. NADH and FAD have distinct absorption and emission spectra (Figure 1.12), and the optical redox ratio, calculated as the ratio of emission intensities of NADH and FAD, serves as a measurement of metabolism alterations in cells [22].

In a Raman scattering spectrum, each peak corresponds to a vibrational mode of a molecule. Therefore each molecule has a very specific Raman fingerprint. Some examples of Raman scattering spectra from biologically relevant compounds are shown in Figure 1.13. For simple systems, Raman scattering provides a mean of compound identification and quantification. On the other hand, Raman scattering spectra of highly complex systems, such as cells and tissues, can be challenging to interpret in an absolute way. However, pathological changes that lead to molecular and structural alterations can be reflected in the Raman spectra and allow Raman scattering to be utilized as a diagnostic tool. Indeed, Raman scattering has been used for this purpose

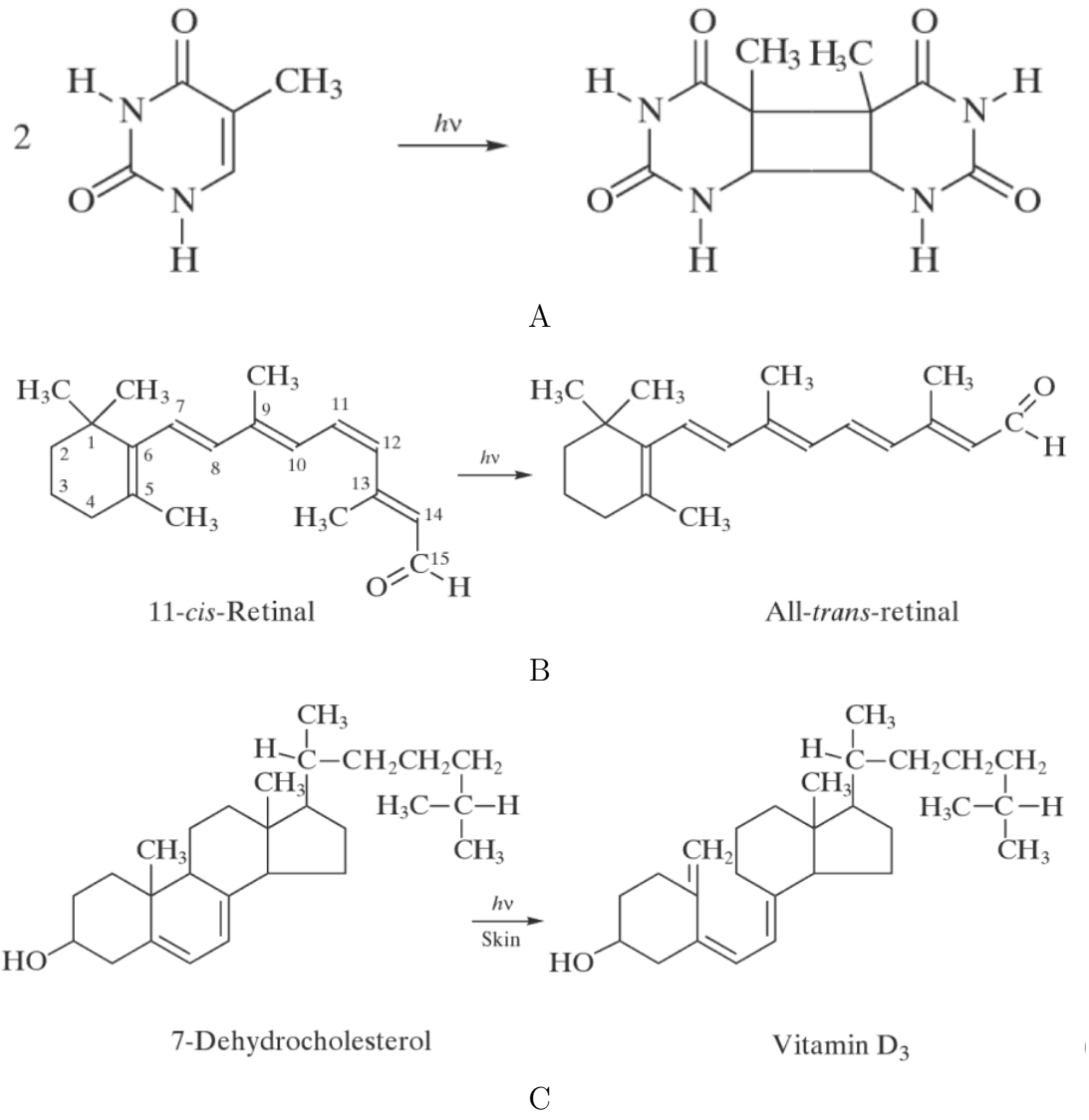
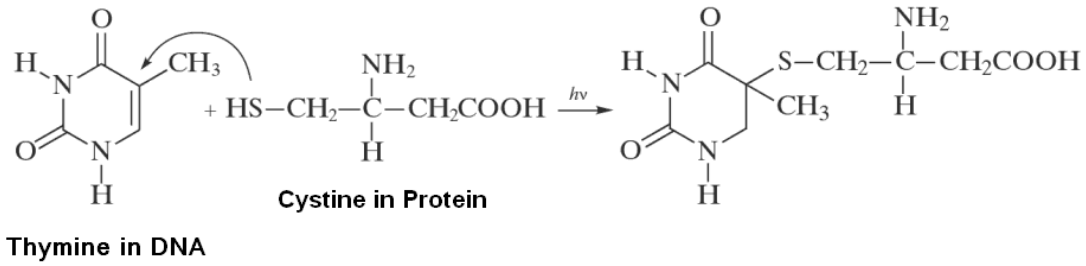
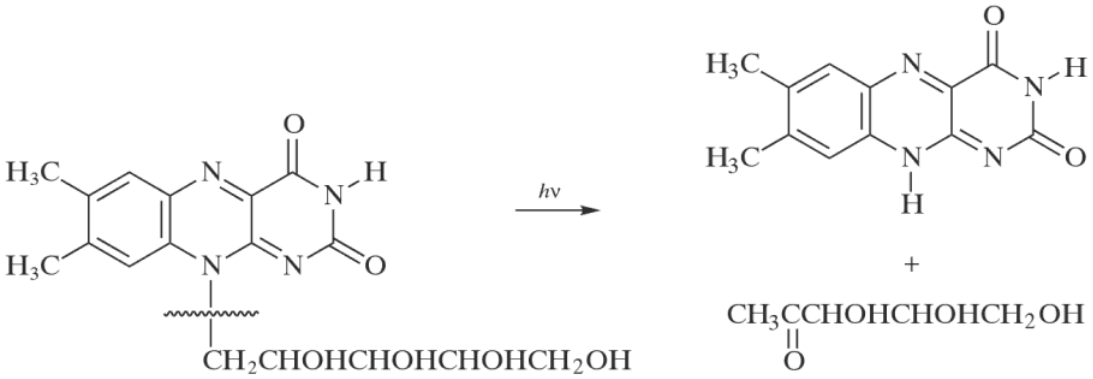


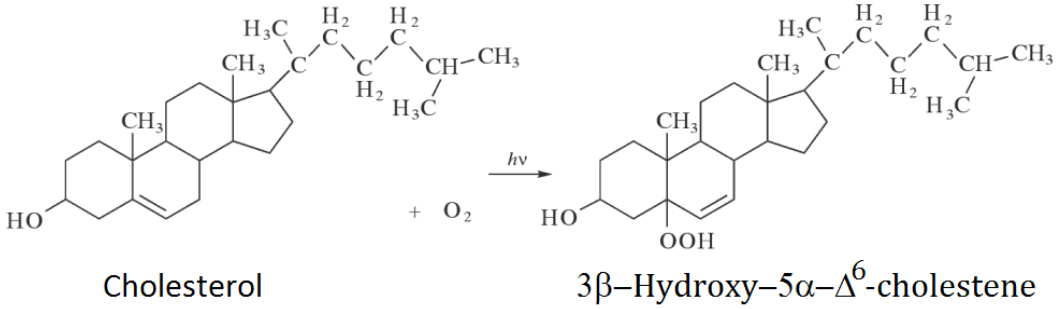
Figure 1.9: Examples of photochemical process in biomolecules **A:** Photodymerization of thymine **B:** photoisomerization of 11-*cis*-retinal **C:** Rearrangement of 7-dehydrocholesterol to vitamin D. Reprinted with permission from [20]. Copyright 2003, John Wiley and Sons.



D

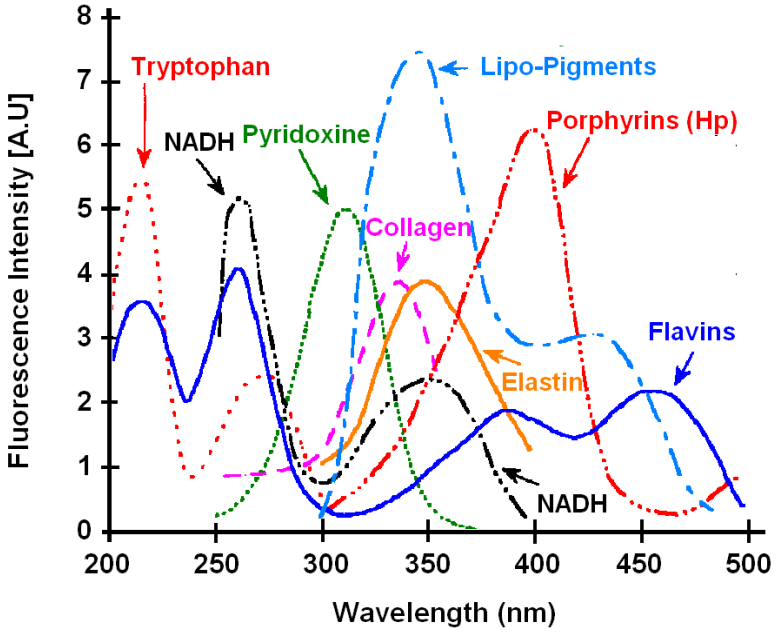


E

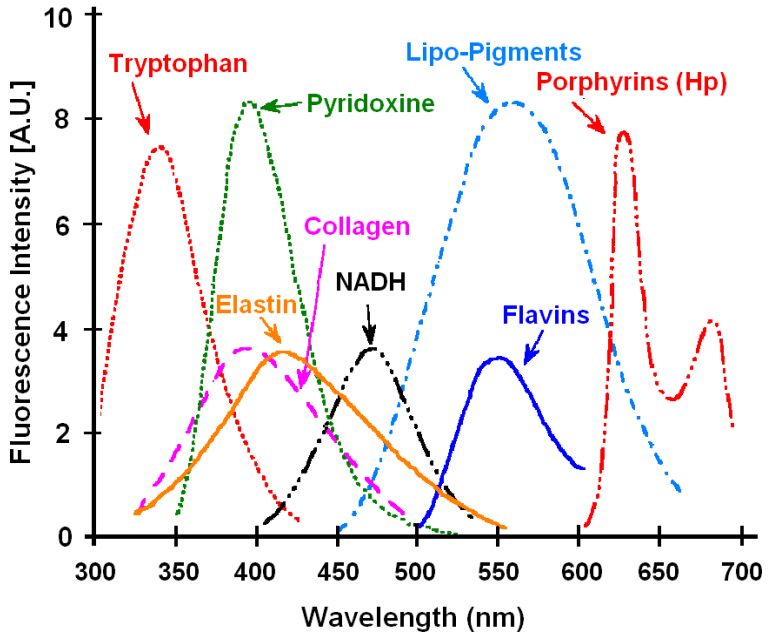


F

Figure 1.10: Examples of photochemical process in biomolecules continued **D:** Light induced protein-DNA crosslinking **E:** Rivoflavin photofragmentation **F:** Photooxidation of cholesterol. Reprinted with permission from [20]. Copyright 2003, John Wiley and Sons.



A



B

Figure 1.11: Spectra of fluorophores present in biological tissue **A:** Absorption and **B:** Emission. Reprinted with permission from [21]. Copyright 2007, Springer.

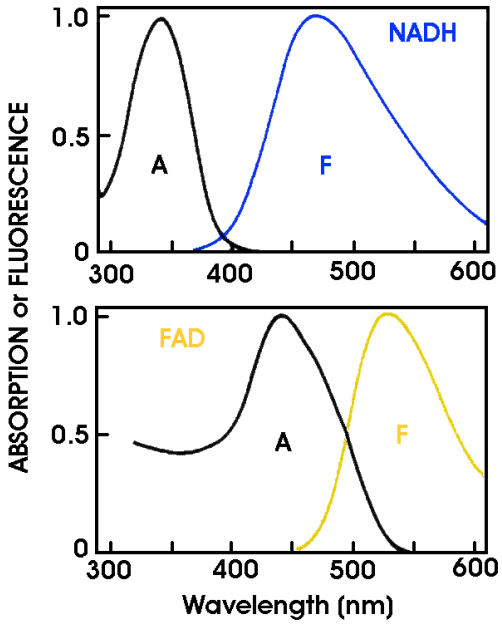


Figure 1.12: Absorption and emission of NADH and FAD, coenzymes involved in cellular metabolism. Reprinted with permission [21]. Copyright 2007, Springer.

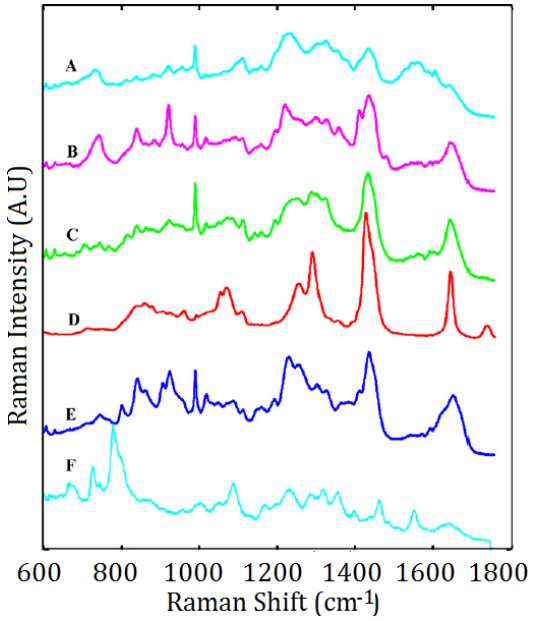


Figure 1.13: Raman spectra of pure compounds. **A:** Blood; **B:** mucus; **C:** proteins; **D:**collagen and **F:** calf thymus DNA. Reprinted with permission from [23]. Copyright 2009, Elsevier.

for a wide range of diseases such as breast cancer [24], lung cancer [25], leukemia [26], prostate and bladder cancer [27], oral carcinoma [28], atherosclerosis [29], bone [30] and macular degeneration [31]. The low concentration of molecules within a cell and the subtle spectral changes make Raman scattering difficult to implement and suggest SERS may be better suited. In addition, SERS is excellent for single molecule inspection, serving as an excellent alternative for protein structure studies [32–35].

In addition to optical responses, tissue and cell also undergo photochemical processes and thermal effects initiated by light absorption. At the tissue level, such processes, going from lower to higher temperature thresholds, are coagulation, vaporization, carbonization and melting. Above 60^oC, the tissue becomes necrotic (i.e., dead), a process known as coagulation. Vaporization is reached when the water within the tissue starts to evaporate (that is at 100^oC). The damage produced by vaporization is purely mechanical caused by the expansion of the water. Carbonization occurs when the tissue reaches over 150^oC, and the organic compounds are converted to carbon. For sufficiently high temperatures tissue can also melt, which is used for tissue welding [20].

1.7 References

[1] D. A. Long, *The Raman effect : A unified treatment of the theory of Raman scattering by molecules* (Wiley, New York, NY, 2002).

[2] K. Meletov, E. Liarokapis, J. Arvanitidis, K. Papagelis, D. Palles, G. Kourouklis, and S. Ves, "On the nature of laser irradiation induced reversible softening of phonon modes in C60 single crystals," Anal Chem Letters **290**, 125–130 (1998).

[3] G. Fini, "Application of Raman spectroscopy pharmacy," J Raman Spectrosc **35**, 335–337 (2004).

[4] M. Moskovits, "Surface-enhanced spectroscopy," Rev Mod Phys **57**, 783–826 (1985).

[5] M. Moskovits, "Surface-enhanced Raman spectroscopy: A brief retrospective," J Raman Spectrosc **36**, 485–496 (2005).

[6] S. Maier, *Plasmonics: Fundamentals and Applications* (Springer, New York, NY, 2007).

[7] P. B. Johnson and R. W. Christy, "Optical Constants of the Noble Metals," Phys Rev B **6**, 4370 (1972).

[8] M. Liu and P. Guyot-Sionnest, "Optical properties of rodlike and bipyramidal gold nanoparticles from three-dimensional computations," Phys Rev B **76**, 235 428 (2007).

[9] E. Fort and S. Grésillon, "Surface enhanced fluorescence," J Phys D Appl Phys **41**, 013 001 (2008).

[10] R. Boyd, *Nonlinear Optics* (Academic Press, 2008).

[11] R. Chang and T. Furtak, *Surface Enhanced Raman Scattering* (Plenum Press, 1982).

[12] A. M. Michaels, J. Jiang, and L. Brus, "Ag nanocrystal junctions as the site for surface-enhanced Raman scattering of single Rhodamine 6G molecules," J Phys Chem B **104**, 11 965–11 971 (2000).

[13] K. A. Bosnick, J. Jiang, and L. E. Brus, "Fluctuations and local symmetry in single-molecule Rhodamine 6G Raman scattering on silver nanocrystal aggregates," J Phys Chem B **106**, 8096–8099 (2002).

[14] P. K. Jain, X. Huang, I. H. El-Sayed, and M. A. El-Sayed, "Review of some interesting surface plasmon resonance-enhanced properties of noble metal nanoparticles and their applications to biosystems," Plasmonics **2**, 107–118 (2007).

[15] D. D. Evanoff Jr and G. Chumanov, "Synthesis and optical properties of silver nanoparticles and arrays," ChemPhysChem **6**, 1221–1231 (2005).

[16] J. J. Mock, M. Barbic, D. R. Smith, D. A. Schultz, and S. Schultz, "Shape effects in plasmon resonance of individual colloidal silver nanoparticles," J Chem Phys **116**, 6755–6759 (2002).

[17] K. Kneipp, H. Kneipp, I. Itzkan, R. R. Dasari, and M. S. Feld, "Surface-enhanced Raman scattering and biophysics," J Phys-Condens Mat **14**, R597–R624 (2002).

[18] C. Douketis, T. L. Haslett, Z. Wang, M. Moskovits, and S. Iannotta, "Self-affine silver films and surface-enhanced Raman scattering: Linking spectroscopy to morphology," J Chem Phys **113**, 11 315–11 323 (2000).

[19] B. Y. S. Kim, J. T. Rutka, and W. C. W. Chan, "Current concepts: Nanomedicine," New Engl J Med **363**, 2434–2443 (2010).

[20] P. Prasad, *Introduction to Biophotonics* (Wiley-Interscience, Hoboken, NJ, 2003).

[21] J. Lakowicz, *Principles of fluorescence spectroscopy, Chapter 3: Fluorophores* (Springer, New York, NY, 2006).

[22] J. H. Ostrander, C. M. McMahon, S. Lem, S. Millon, J. Brown, V. L. Seewaldt, and N. Ramanujam, "Optical redox ratio differentiates breast cancer cell lines based on estrogen receptor status," Cancer Res **70**, 4759–4766 (2010).

[23] A. Beljebbar, O. Bouché, M. Diébold, P. Guillou, J. Palot, D. Eudes, and M. Manfait, "Identification of Raman spectroscopic markers for the characterization of normal and adenocarcinomatous colonic tissues," Crit Rev Oncol Hemat **72**, 255 – 264 (2009).

[24] A. S. Haka, K. E. Shafer-Peltier, M. Fitzmaurice, J. Crowe, R. R. Dasari, and M. S. Feld, "Diagnosing breast cancer by using Raman spectroscopy," P Natl Acad Sci USA **102**, 12 371–12 376 (2005).

[25] N. D. Magee, J. R. Beattie, C. Carland, R. Davis, K. McManus, I. Bradbury, D. A. Fennell, P. W. Hamilton, M. Ennis, J. J. McGarvey, and J. S. Elborn, "Raman microscopy in the diagnosis and prognosis of surgically resected nonsmall cell lung cancer," J Biomed Opt **15**, 026 015 (2010).

[26] J. C. Martínez-Espinosa, J. L. González-Solís, M. L. Miranda-Beltrán, C. Soria-Fregoso, J. Medina-Valtierra, and C. Frausto-Reyes, "Detection of leukemia with blood samples using raman spectroscopy and multivariate analysis," **1142**, 99–103 (2009).

[27] T. J. Harvey, E. C. Faria, A. Henderson, E. Gazi, A. D. Ward, N. W. Clarke, M. D. Brown, R. D. Snook, and P. Gardner, "Spectral discrimination of live prostate and bladder cancer cell lines using Raman optical tweezers," J Biomed Opt **13** (2008).

[28] R. Malini, K. Venkatakrishna, J. Kurien, K. M. Pai, L. Rao, V. B. Kartha, and C. M. Krishna, "Discrimination of normal, inflammatory, premalignant, and malignant oral tissue: A Raman spectroscopy study," Biopolymers **81**, 179–193 (2006).

[29] R. Rocha, L. Silveira Jr, A. B. Villaverde, C. A. Pasqualucci, M. S. Costa, A. Brugnera Jr, and M. T. T. Pacheco, "Use of near-infrared Raman spectroscopy for identification of atherosclerotic plaques in the carotid artery," Photomed Laser Surg **25**, 482–486 (2007).

[30] J. Nyman, A. Makowski, C. Patil, T. Masui, E. OQuinn, X. Bi, S. Guelcher, D. Nicollela, and A. Mahadevan-Jansen, "Measuring differences in compositional properties of bone tissue by confocal Raman spectroscopy," Calcif Tissue Int **89**, 111–122 (2011).

[31] I. V. Ermakov, M. Sharifzadeh, M. Ermakova, and W. Gellermann, "Resonance Raman detection of carotenoid antioxidants in living human tissue," J Biomed Opt **10**, 11791181 (2005).

[32] M. Feng and H. Tachikawa, "Surface-enhanced resonance Raman spectroscopic characterization of the protein native structure," J Am Chem Soc **130**, 7443–7448 (2008).

[33] F. Wei, D. Zhang, N. J. Halas, and J. D. Hartgerink, "Aromatic amino acids providing characteristic motifs in the raman and SERS spectroscopy of peptides," J Phys Chem B **112**, 9158–9164 (2008).

[34] X. X. Han, Y. Kitahama, T. Itoh, C. X. Wang, B. Zhao, and Y. Ozaki, "Protein-mediated sandwich strategy for surface-enhanced raman scattering: Application to versatile protein detection," Anal Chem **81**, 3350–3355 (2009).

[35] G. Das, F. Gentile, M. L. Coluccio, A. M. Perri, A. Nicastri, F. Mecarini, G. Cojoc, P. Candeloro, C. Liberale, F. De Angelis, and E. Di Fabrizio, "Principal component analysis based methodology to distinguish protein SERS spectra," J Mol Struct (2011), article in Press.

CHAPTER 2

PATHOGENESIS AND CURRENT DIAGNOSTIC TECHNIQUES

2.1 Cancer

There are more than 100 distinct types of cancer [1]. Hanahan and Weinberg suggested in 2000 that all the cancerous genotypes are the manifestation of six essential alterations in the cell: self-suffiency in cell growth, insensitivity to anti-growth signals, evasion of apoptosis, limitless replicative potential, sustained angiogenesis and tissue invasion and metastasis. Recently it has been suggested that metabolomic alterations have to be included in that list [2]. In the present chapter, the mechanisms accountable for these features are described.

2.1.1 Pathogenesis: cancer is a genetic disorder

The way cancer pathogenesis is understood has changed radically since the discovery of **oncogenes** and **tumor suppressor genes** (TSG), which are sequences in the DNA that upon activation and deactivation, respectively, could induce cancerous transformation. Although simple in concept, activation of a single oncogene is not enough to induce transformation. Two or more active oncogenes are required, and as transformation progresses cells acquire more mutations [3]. A transformation based solely on activation of oncogenes will take longer than what is observed clinically, as the chance for each mutation to happen is very low. TSGs are genes that help regulate responses in normal cells aimed to control growth and proliferation, and cancerous cells appear as if they have lost the normal expression of these genes. Still, the chances of DNA mutations at the right places are very low, especially taking into account that each gene has two copies in each cell. Unfortunately, there is an additional mechanism by which cancerous cells can overcome such 'problems'. **Loss**

of heterozygosity (LOH) is the process by which malignant cells are able to discard the chromosomal arm that has the intact copy of the TSG and replace it with a copy of the arm that has the mutated version [3].

The chances of a gene mutation are not only stochastically low, but are further reduced by the existence of molecular machinery that is capable of detection and repair of defective DNA. Faulty DNA repair, in turn, accelerates transformation and seems to be a feature present in a great majority of human cancers [3]. So far, mechanisms that affect the DNA sequence directly have been described. Unfortunately, additional mechanisms (i.e., epigenetic mechanisms) that assist cancerous transformation exits, such as the case of **DNA methylation**, which is an important mechanism to regulate gene transcription. It results from the addition of a methyl group to a cytosine ring. The methylation pattern is inherited from parents and it is maintained by the methylase. The methyl group regulates expression by creating a physical barrier between the DNA and the transcription machinery, and modifying histones, thus helping to create large parts of inactive chromatin. Hypermethylation, the abnormal addition of methyl groups, can lead to the deactivation of TSG whereas hypomethylation, the abnormal removal of methyl groups, can lead to the activation of oncogenes [4].

Cancer cells can grow and divide indefinitely, while normal cells only divide for a finite number of generations. In a normal cell, the telomeres of the chromosome shorten during each cycle of growth and division. Eventually they become too short to perform their job of preventing fusions among different chromosomes. At this point the cell enters the death program and no further division takes place. Cancerous cells seem to have activated the expression of telomerase, the enzyme that repairs telomeres. Expression of telomerase is detectable in almost all human cancers [3]. In many cases, collaborative genetical and epigenetical modifications give cancer cells the acquired ability to migrate and establish themselves elsewhere in the body. This process is dubbed **invasion and metastasis**.

2.2 Diagnosis

Histopathology is the gold standard for cancer diagnosis. It bases the diagnosis on structural changes of the tissue. Paraffin embedded sections of the tissue are stained with hematoxylin and eosin (H&E) and observed under an optical microscope (See Figure 2.1). The result strongly depends on proper handling of the tissue sample. Moreover, the sample must undergo appropriate preparation and processing, which could take up several days [5]. Even if ideal conditions have been met up to this point, uncertainty in the interpretation of biopsies is an inherent problem of histopathological diagnosis because i) the sample may have been obtained at a stage when the cancerous features have not yet been fully developed, ii) treatment procedures, like irradiation and chemotherapy can change the morphology of the tissue, iii) mechanical, electrical and thermal stress can alter morphological features and iv) some malignancies, by nature, differ very little from other malignant, benign or reactive disorders [5].

In addition to histopathology, cytopathology, the study of the morphology of individual cells, is often used as a screening technique. Cytopathology has high sensitivity (i.e., few cancer cases are missed) but has low specificity, meaning that patients without cancer will undergo additional testing, increasing the overall cost

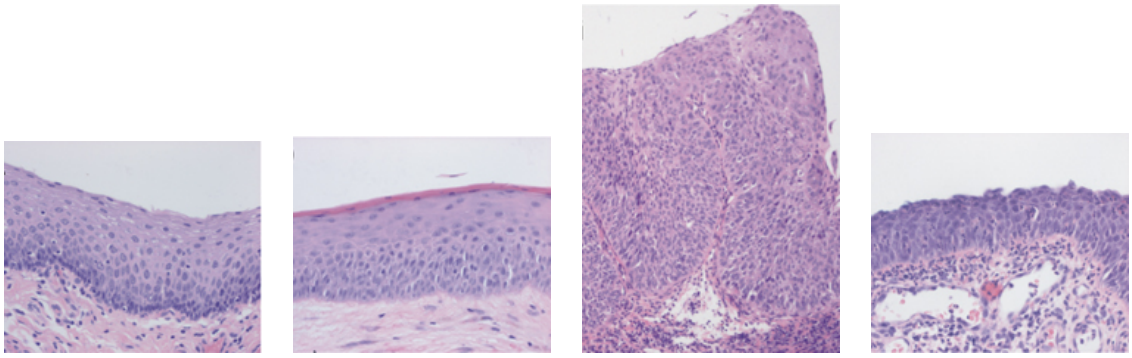


Figure 2.1: H&E stained images of **A:** Normal squamous tissue, B: abnormal hyperplasia, C: carcinoma in situ and D: dysplasia. All images are X200 magnification. Reprinted with permission from [6]. Copyright 2003, Nature Publishing Group.

of diagnosis [5]. An important problem in diagnosis is the interobserver variability, which has been estimated to be as high as 10% [7]. Additionally, discrepancies in the diagnosis made from histopathology and cytopathology within the same institution can be as high as 12% [8].

In approximately 10% of all oncology cases, histopathology and cytopathology are not sufficient to confirm a diagnosis. In such cases, additional, more specialized and more expensive tests, like histochemical stains, electron microscopy and immunochemistry, are used. Flow-cytometry is another important technique used to identify different cell populations within a sample which is done by labeling cells with fluorescent labels that bind to specific membrane markers. Finally, techniques aimed at measuring gene expression and DNA mutations are also used, although they are not widely available. Examples of such techniques are polymerase chain reaction and in situ hybridization and fluorescent in situ hybridization [5].

All things considered, novel diagnostic tools are a necessity for a faster, cheaper and more accurate diagnosis. Any new technique will have to fulfil many requirements. i)The new processes should be simple to implement, with mandatory feasibility for automatization, which in turn will lead to faster diagnosis and lower interobserver variability, ii) the new techniques should be label/reagent free, which will help reduce costs, and importantly, iii) the new methods should have not only high sensitivity but also high specificity, thus reducing the necessity for additional tests. Simultaneously fulfilling these three requirements is the challenge of the new generation of diagnostic tools.

2.3 Cellular metabolism in normal and neoplastic cells

Cells obtain the energy necessary for their survival and proliferation via two mechanisms, Glycolysis and the Krebs cycle. Under normal conditions (i.e., normal oxygen tension), cells convert one molecule of glucose into 36 adenosine triphosphate [1]

[1]Adenosine triphosphate is a nucleotide with three phosphate groups. Energy is stored in ATP molecules in the phosphate bond which can be released by the hydrolysis of ATP into ADP (Adenosine diphosphate).

(ATP) molecules through the process called the Krebs cycle or the tricarboxylic acid (TCA) cycle. In conditions with limited oxygen (hypoxic conditions), cells switch their metabolism mechanism to anaerobic glycolysis which results in the increased production of lactic acid.

Some cancer cells show a steep increase in the consumption of glucose and produce lactic acid instead of metabolizing glucose through the TCA cycle, even in the presence of sufficient oxygen. This phenomenon is termed aerobic glycolysis or Warburg effect, after Otto Warburg who received the Nobel prize in 1931 for his work on cellular metabolism.

In this section a brief introduction to glycolysis, the tricarboxylic acid cycle and the Warburg effect are presented to give context to the work developed in Chapter 4.

2.3.1 Glycolysis

Glycolysis is the metabolic pathway in which glucose is broken and oxidized to form pyruvate. The main role of glycolysis is the formation of ATP and nicotinamide adenine dinucleotide (NAD), although glycolysis also provides biosynthetic precursors for the synthesis of fatty acids, amino acids and nucleotides.

The glycolytic pathway cleaves and oxidizes 1 molecule of glucose to 2 molecules of each pyruvate and ATP, and reduces 2 molecules of NAD in a 10 step process summarized in Figure 2.2. Each step involves an enzyme that catalyzes the reaction. Few of these steps require energy to proceed but there is at the end a net production of ATP molecules. Pyruvate, the other molecular product of glycolysis, is the starting point for the TCA cycle, as described in the next section. NAD is an important coenzyme involved in reduction-oxidation (REDOX) reactions as it serves as an electron carrier. During glycolysis, molecules of NAD are reduced to NADH, a fact that as explained below, can be used to spectroscopically measure alterations to cellular metabolism.

2.3.2 Tricarboxylic acid cycle

For the TCA cycle to take place, a preparatory step is required: Pyruvate needs to be oxidized into Acetyl-Co, through a process called pyruvate decarboxylation.

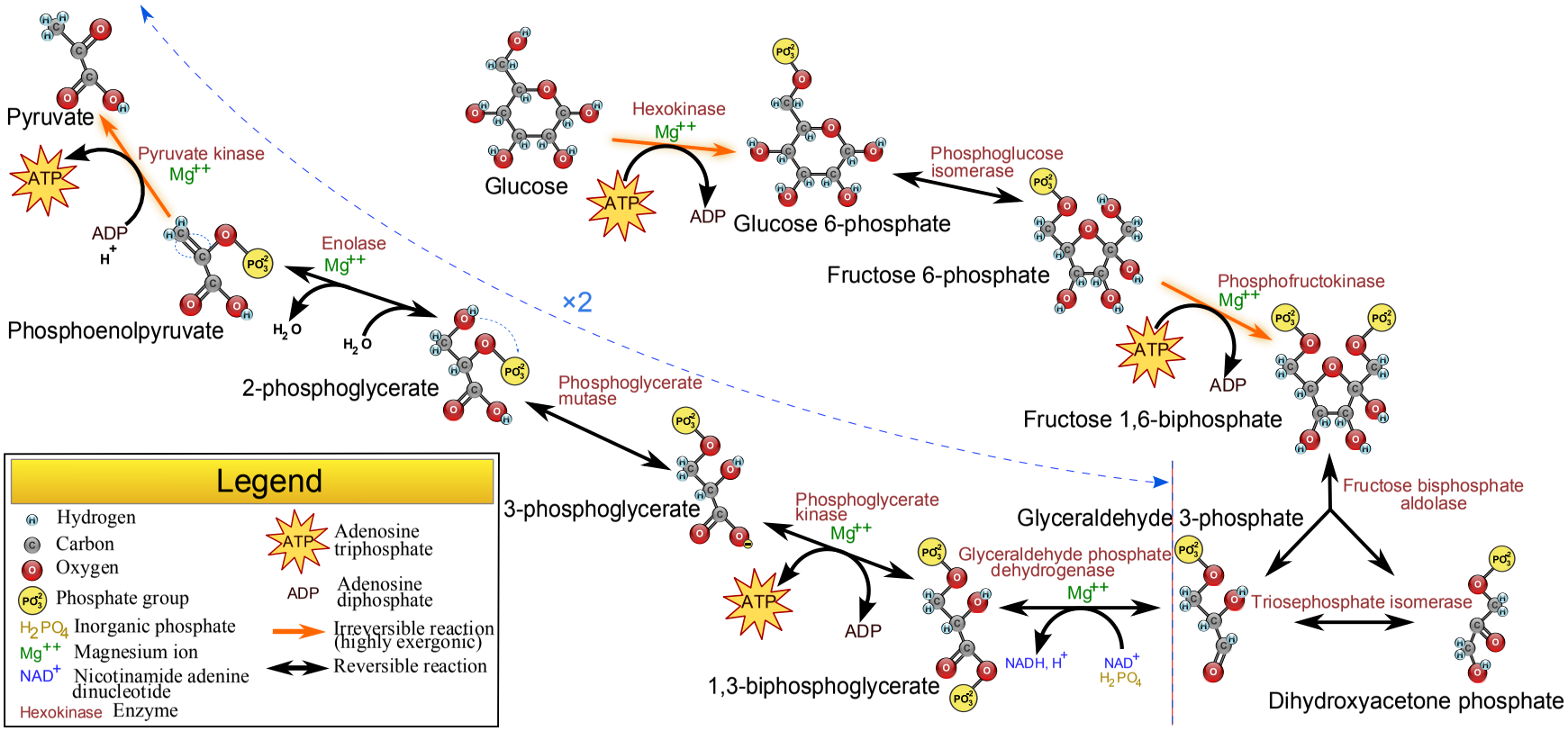


Figure 2.2: Glycolysis pathway. Text in green denotes enzyme. Green Boxes are assigned to reactions that take place twice. Reprinted under the *Creative Commons Attribution-Share Alike 3.0 Unported license* from [9].

Unlike glycolysis that takes place in the cytosol, pyruvate decarboxylation and the TCA cycle take place in the mitochondria. As shown in Figure 2.3, Acetyl-Co is the true starting point of the TCA cycle, which is converted into citrate, and at the end of the cycle, the oxaloacetate is recovered and the cycle starts again. The TCA cycle takes electrons from Acetyl-Co and uses them to form NADH and to reduce Flavin Adenosine (FAD) to form $FADH_2$. Oxygen is necessary at the end of the cycle to regenerate NAD and FAD. The net products of the TCA cycle are two ATP molecules. So, where are the additional 34 ATP molecules coming from? The answer is oxidative phosphorylation.

Oxidative phosphorylation is the final part of the aerobic metabolism, and it uses the electrons in NADH and $FADH_2$ to generate a proton gradient across the mitochondrial membrane which then activates ATP Synthase, the protein complex that synthesizes ATP. Redox reactions are done by four protein complexes (including ATP Synthase) located in the inner membrane of the mitochondria in the presence of oxygen. Oxygen is the last electron acceptor and is essential for the regeneration of NAD and FAD. From the oxidative phosphorylation between 30 and 36 ATP molecules are produced.

2.3.3 Hypoxia and Warburg effect

Warburg effect, or aerobic glycolysis, is the name given to the observed effect in which cancerous cells display a higher glucose uptake and a higher lactic acid production even under normal oxygen (normoxic) conditions. For reasons that will be described, cancerous cells do not metabolize glucose through the TCA cycle but instead they do so by just glycolysis, which in turn increases the production of pyruvate. Glycolysis involves the reduction of NAD to NAHD. In a normal cell under sufficient oxygen conditions pyruvate will enter the TCA cycle. However, when there is not enough oxygen available to proceed with the oxidative phosphorylation and therefore oxidize NADH back to NAD, glycolysis is not possible due to a shortage in NAD. Instead, pyruvate can be used as a substrate for the oxidation of NADH in which, by the catalysis of lactate dehydrogenase (LDH), pyruvate is converted into

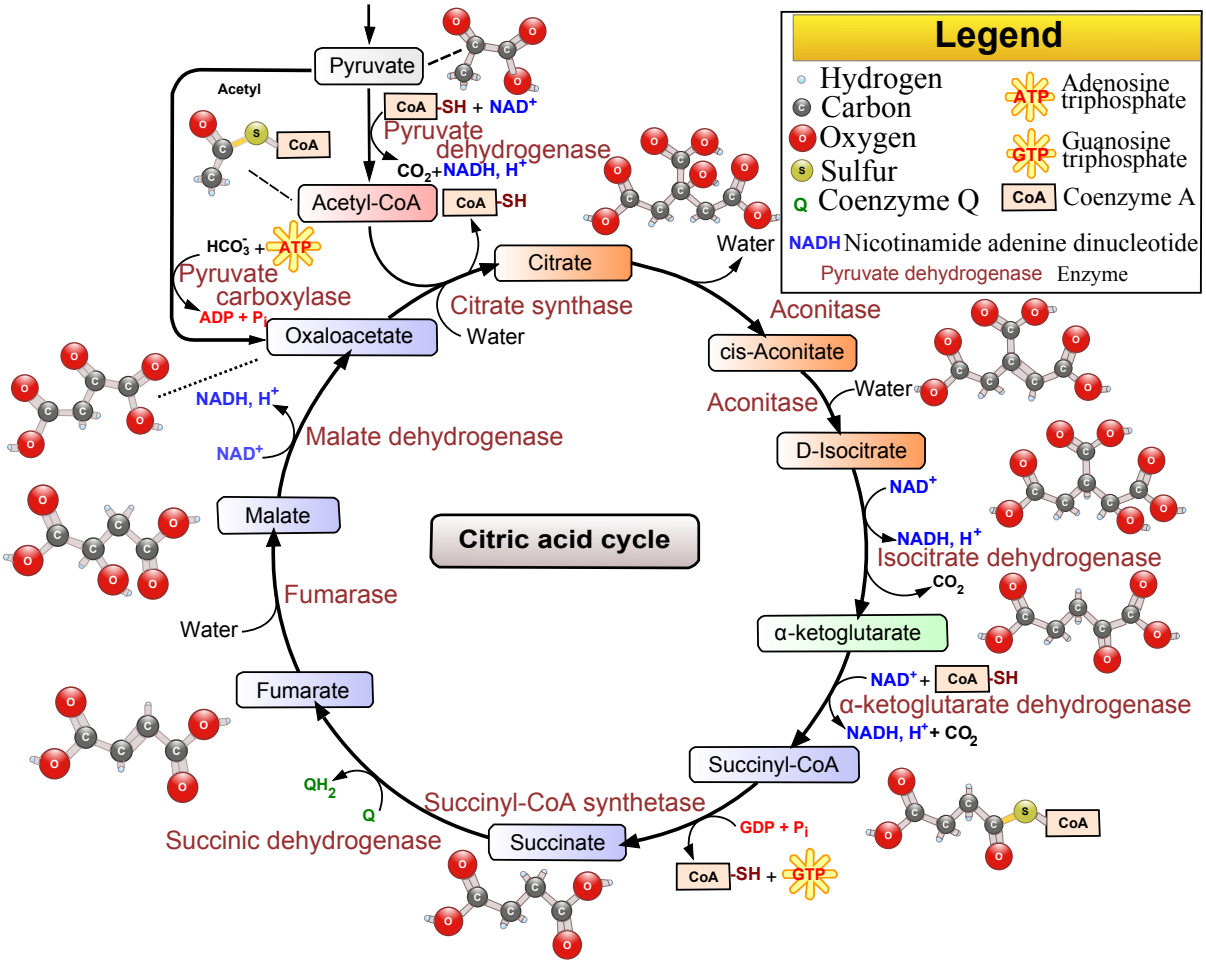


Figure 2.3: Tricarboxylic acid cycle. Reprinted under the *Creative Commons Attribution-Share Alike 3.0 Unported license* from [10].

lactic acid.

One of the ways that normal cells respond to hypoxia is through hypoxia inducible factor 1 (HIF-1). HIF-1 is a transcription factor[2] that in the presence of oxygen, gets ubiquitylated and thus tagged for proteasome degradation. Under hypoxic conditions HIF-1 activates the expression of genes involved in angiogenesis, metabolism, cell proliferation, among others [11–13] (See Figure 2.4). In cancerous cells, many mutations have been identified that lead to increased HIF-1 activity [13,14], which in turn increases glucose uptake and pyruvate conversion to lactate as well as reduces mitochondrial activity and mass [14–16].

Several HIF-1 independent mutations have been implicated in the Warburg effect [16]. Such is the case of the AKT gene which, in turn, also promotes an increase in glucose uptake and aerobic glycolysis. Myc genes code for transcription factors that regulate growth and cell cycle entry [3] and are activated in a broad range of cancers. Myc genes have been associated with both activation of anaerobic glycolysis and increased mitochondrial activity. Other genes involved in Warburg effect are the succinate dehydrogenase (SDH) and fumarate hydratase (FH) [16].

Metabolic transformation is important for cancer cells as it not only provides enough energy, but it provides many of the building blocks necessary for cell proliferation. Additionally, aerobic glycolysis has been linked to cellular immortalization and antiapoptotic effects [16].

2.4 Myeloproliferative neoplasms

Myeloproliferative neoplasms (MPNs) belong to a family of diseases that originate in a somatic mutation[3] in a self-renewing hematopoietic stem cell. MPNs include four main diseases: chronic myelogenous leukemia (CML), polycythemia vera (PV), essential thrombocythemia (ET), and primary myelofibrosis (PMF). CML is characterized

[2]A protein that can increase (upregulate) or suppress (downregulate) the expression of genes.

[3]A somatic mutation is an acquired mutation as opposed to a germline mutation, which is an inherited mutation.

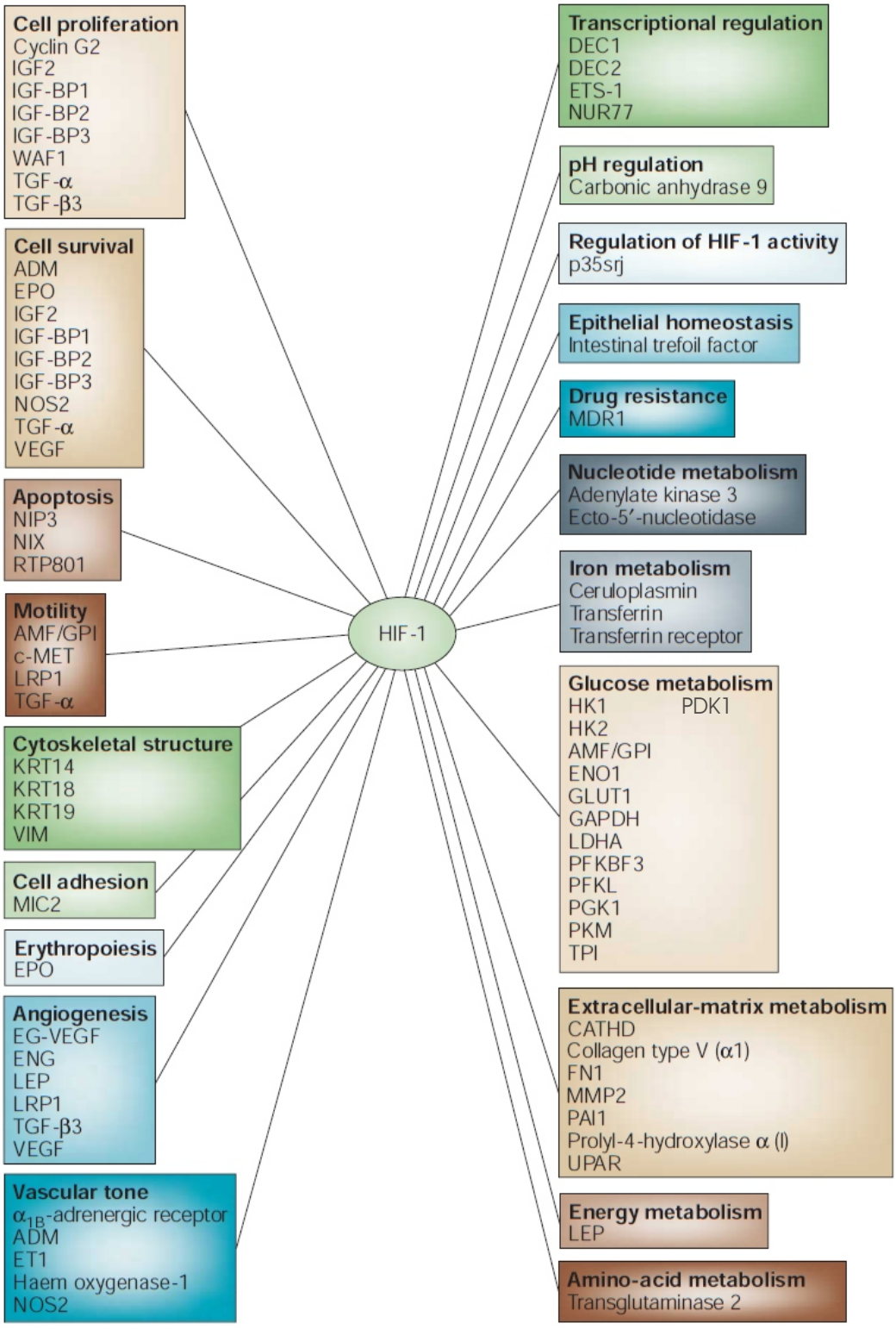


Figure 2.4: Genes that are involved in many processes are transcriptionally activated by HIF-1. Reprinted with permission from [13]. Copyright 2003, Nature Publishing Group.

by the BCL-BCR fusion gene (i.e., Philadelphia chromosomal abnormality). PV, ET and PMF are Philadelphia chromosome negative The major features of MPNs are increased production of red blood cells in PV, high platelet count in ET and bone marrow fibrosis in PMF [17–19]; all three share common characteristics such as bone marrow hypercellularity, predisposition to thrombosis and hemorrhage and increased risk of transformation to acute myeloid leukemia [18].

Nonphiladelphia MPNs are thought to originate in the transformation of a hematopoietic stem cell, but their molecular pathogenesis has yet to be completely elucidated [18, 19]. Nonetheless, key molecular features of MPNs have been discovered in recent years and are briefly outlined in the following section.

2.4.1 Pathogenesis

In 2005 the valine to phenylalanine substitution at the 617 residue of the tyrosine kinase gene that encodes for *JAK2* was reported by several groups [18, 20–23]. The *JAK2* V617F mutation is not in the germline, but it is rather acquired, and it has been shown to be present in 90-95% of PV patients as well as in 50-60% of ET and PMF patients [17, 19]. This mutation causes cytokine[4]-independent activation of JAK-STAT, PI3K, and AKT pathways and mitogen-activated protein kinase (MAPK) and extracellular signal-regulated kinase (ERK), all of which are linked to the erythropoietin receptor (EPOR) that has an important role in cellular auto-inhibition, in other words, apoptosis, proliferation and differentiation [19]. The mutation is at a motif that renders EPOR ubiquitously active, even in the absence of erythropoietin, the cytokine produced in the kidney and liver responsible for EPOR activation under normal conditions (see Figure 2.5). The *JAK2* V617F mutation explains the formation of erythropoietin independent erythroid colonies (i.e., bone marrow progenitor cells colonies), a hallmark of polycythemia vera , but is not necessary as demonstrated by the formation of colonies in PV patients without the *JAK2* V617F mutation [24].

[4]Small proteins used for intercellular signalling. Citokines trigger cellular responses by binding to receptors located at the surface of the cells.

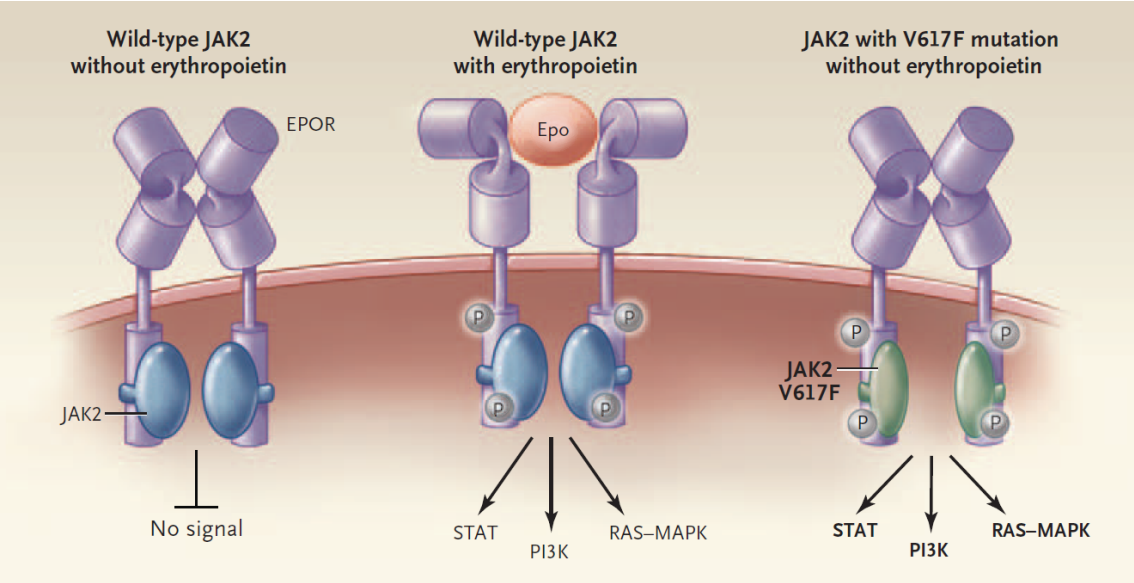


Figure 2.5: Under normal conditions, EPOR is inactive in the absence of erythropoietin. The binding of erythropoietin induces conformational changes to EPOR, causing the phosphorylation of JAK2 and thus the activation of signalling pathways. In cells with the V617F mutation the binding of erythropoietin is outwitted. Reprinted with permission from [19]. Copyright 2006, Massachusetts Medical Society.

Although the discovery of the *JAK2* V617F mutation helps explain many of the key features of PV, there is strong evidence that *JAK2* V617F mutation is not the initiating mutation. For example, it is known from *in vitro* transfection experiments that wild *JAK2* is dominant over *JAK2* V617F, and it is also known that most PV patients are heterozygous for the *JAK2* V617F mutation. Additionally, in some women, the number of clonal granulocytes outnumber those of *JAK2* V617F positive granulocytes [19]. Finally, the identification of a patient whose complete population of granulocytes showed the deletion of chromosome 20q but only 7% *JAK2* mutant allele frequency [24]. All suggesting that the *JAK2* V617F mutation is not enough to produce the polycythemia vera phenotype.

Recently it has been found that TET2 (TET oncogene family 2) and CBL mutations are involved in the pathogenesis of MPNs [25, 26], but there is evidence that neither TET2 [27] nor CBL [26] are disease-initiating mutations.

2.4.2 Diagnosis

The World Health Organization (WHO) classifies ET, PV, PMF, CML and the less frequent chronic neutrophilic leukemia, chronic eosinophilic leukemia/hypereosinophilic syndrome, systemic mastocytosis, and unclassifiable MPD as Myeloproliferative Neoplasms. As describe above, ET, PV and PMF share a lot in common, and clinically, they mimic each other and some other malignant and nonmalignant disorders. Although the *JAK2* V617F mutation was a major advance in the diagnostics of MPNs, it is not specific, meaning it cannot be used to distinguish between one MPN from another. On the other hand, the absence of the *JAK2* V617F mutation has little diagnostic value in ET and PMF, since only about 50% of patients will have the mutation [28], and it does not exclude PV. Many tests are used for the diagnosis of MPN to differentiate among them and to assure diagnosis in the case of *JAK2* V617F negative result.

The differential diagnosis of PV usually starts when an increased red cell mass (RCM) is observed. Increased RCM can be due to several reasons factors: reactive erythrocytosis, drugs, pulmonary disease, high altitude, etc. It is not straightforward but absolutely necessary to rule out those situations. Measurements of hematocrit[5] and hemoglobin levels, although routinely done, do not always allow distinction among MPNs. Table 2.1 summarizes additional tests used for the differential diagnosis of PV.

PMF is characterized by bone marrow fibrosis and hypercellularity. Its diagnosis relies on the right identification of morphologically atypical megakaryocytes, megakaryocytes proliferation and reticulin and/or collagen fibrosis.

Essential thrombocythemia is characterized by an elevated platelet count (i.e., thrombocytosis). Thrombocytosis can be caused by a number of reasons: infection, renal disorders, blood loss, tissue damage, chronic inflammation, among others [29]. The first step is to exclude the possibility of reactive thrombocytosis due to the reasons mentioned. The second step is to distinguish ET from any other myeloid disorder, including CML. At this point, diagnosis of ET becomes challenging as one

[5]Percentage of blood volume occupied by red blood cells.

Table 2.1: Additional laboratory tests used in the differential diagnosis of polycythemia vera [29].

Test	Technique	Description
EpO Level	ELISA	An elevated EpO level rules out PV
Endogenous erythroid colonies (EEC)	Cell Colonies	EEC is a feature of PV, but it is not standardized, not available clinically and negative test does not exclude PV [29].
Bone marrow biopsy	Morphology	Will not distinguish PV from PMF or ET.
Splenic enlargement	CT or ultrasound	

of the major criteria for ET diagnosis is that it does meet the criteria for CML, PV or PMF, i.e., ET is diagnosed by exclusion. Table 2.2 summarizes the WHO diagnostic criteria for myeloproliferative neoplasms reported in 2008 [28].

2.5 Hemoglobinopathies

Hemoglobinopathies are a group of disorders that affect red blood cells with abnormal hemoglobin (Hb). They can be divided into expression abnormalities, called *Thalassemias*, and structural disorders. Structural disorders are the result of aberrant hemoglobin molecules, i.e., hemoglobin variants. There are 1118 hemoglobin variants identified [30]; they are due to a point mutation in a globin gene which produces a molecule of hemoglobin that has a single amino acid substitution. They receive the name from where they were discovered, although the first ones were given letters from the alphabet. The most severe symptomatic hemoglobinopathies are the homozygous ones; heterozygous ones, also refered to as traits, tend to be milder or asymptomatic, but some have moderate clinical implications, especially traits of Hb S, the hemoglobin variant associated with sickle cell disease. Table 2.3 lists some relevant Hb variants. For a full updated list see [30].

Many of these variants do not have relevant clinical implications, but some like Hb S and Hb C do. Additionally, variants without clinical manifestations could

Table 2.2: Diagnosis of polycythemia vera (PV) requires meeting either both major criteria and one minor criterion or the first major criterion and 2 minor criteria. Diagnosis of essential thrombocythemia requires meeting all four major criteria. Diagnosis of primary myelofibrosis (PMF) requires meeting all three major criteria and two minor criteria. Modified from [28].

		Polycythemia Vera		**Essential Thrombocythemia**		**Primary Myelofibrosis**
Major Criteria	1	Increased hemoglobin level, hematocrit or red cell mass	1	Increased platelet count	1	Megakaryocyte proliferation and atypical morphology. Reticulin and/or collagen fibrosis. Bone marrow hypercellularity.
	2	*JAK2* V617F Positive	2	Megakaryocyte proliferation with atypical morphology. No or little granulocyte or erythroid Proliferation.	2	Not meeting WHO criteria for CML, PV, MDS, or other myeloid neoplasm.
			3	Not meeting WHO criteria for CML, PV, PMF, MDS or other myeloid neoplasm	3	*JAK2* V617F Positive or no evidence of reactive marrow fibrosis
			4	*JAK2* V617F Positive or no evidence of reactive thrombocytosis		
Minor Criteria	1	Subnormal EpO Level			1	Marked number of immature erythroid and myeloid cells in the circulation
	2	EEC growth			2	Increased serum lactate dehydrogenase
	3	Bone Marrow			3	Anemia
					4	Palpable splenic enlargment

Table 2.3: Representative hemoglobin variants. For a full updated list see [30].

Hb Variant	**Clinical Features**	**Diagnostic Challenges**
S	Sickle cell disease	None
C	Splenic enlargement, mild hemolytic anemia	None
E	Growth retardation, mongoloid faces,moderate splenomegaly,iron overload and mild hemolytic anemia	Hard to differentiate form Hb O
D	Asymptomatic when heterozygous	Hard to differentiate form Hb G
A_2	Elevated in people with β-thalassemia	Hard to differentiate for Hb E and Hb Lepore
F	Asymptomatic	None
O	Asymptomatic when heterozygous	Hard to differentiate form Hb E
G	Asymptomatic when heterozygous	Hard to differentiate form Hb D
Lepore	Asymptomatic	Eludes both electrophoretic and chromatographic methods

cause substantial overestimation of HbA1c, the important marker for diabetes mellitus [31]. Hb variants are commonly identified using electrophoretic or chromatographic techniques. Electrophoresis, at alkaline or acidic pH, is able to resolve common variant such as S, F, A and C but fails to differentiate E from O and D from G. Another electrophoretic technique is Isoelectric Focusing, which despite having better resolution, does not have quantitative capabilities and is very labor intensive [32]. On the other hand, high performance liquid chromatography (HPLC) provides a fast diagnosis but lacks resolution to differentiate common variants such as E and Lepore from A_2 and D from G. Neither electrophoretic nor HPLC has enough sensitivity to detect Hb at low concentrations [32].

Cation exchange HPLC has a better resolution than normal HPLC but suffers from a disadvantage shared with electrophoretic techniques: Since the principle of separation is based on differences in affinity and charge distribution, variants that are neutral elude this methods [31]. Mass spectroscopy (MS) has been used for Hb variant diagnosis as well, but still suffers from some important drawbacks: Due to low mass resolution MS is unable to differentiate among variants with few mass differences such as C, D, and E. Moreover, MS is a qualitative technique, so quantification of important Hb fractions like HbA1c and HbA2 is not possible [31].

2.6 References

[1] D. Hanahan and R. A. Weinberg, "The hallmarks of cancer," Cell **100**, 57–70 (2000).

[2] R. A. Cairns, I. S. Harris, and T. W. Mak, "Regulation of cancer cell metabolism," Nat Rev Cancer **11**, 85–95 (2011).

[3] J. Mendelsohn, *The Molecular Basis of Cancer* (Saunders/Elsevier, Philadelphia, PA, 2008).

[4] P. M. Das and R. Singal, "DNA Methylation and Cancer," J Clin Oncol **22**, 4632–4642 (2004).

[5] R. Lenhard, R. Osteen, T. Gansler, and A. C. Society, *Clinical Oncology* (American Cancer Society, Atlanta, GA, 2001).

[6] P. Chatrath, I. S. Scott, L. S. Morris, R. J. Davies, S. M. Rushbrook, K. Bird, S. L. Vowler, J. W. Grant, I. T. Saeed, D. Howard, R. A. Laskey, and N. Coleman,

"Aberrant expression of minichromosome maintenance protein-2 and Ki67 in laryngeal squamous epithelial lesions," Br J Cancer **89**, 1048–1054 (2003).

[7] M. Markman, "Diagnostic pathology: Challenges and increasing complexity," Oncology **79**, 159–160 (2010).

[8] S. S. Raab and D. M. Grzybicki, "Quality in cancer diagnosis," CA Cancer J Clin **60**, 139–165 (2010).

[9] YassineMrabet, "Glycolysis.svg," (Access July 22 2011), `http://commons.wikimedia.org/wiki/File:Glycolysis.svg`.

[10] wadester16, "Citric_acid_cycle_with_aconitate_2.svg," (Access July 22 2011), `http://commons.wikimedia.org/wiki/File:Citric_acid_cycle_with_aconitate_2.svg`.

[11] G. L. Semenza, "Signal transduction to hypoxia-inducible factor 1," Biochem Pharmacol **64**, 993–998 (2002).

[12] A. L. Harris, "Hypoxia - A key regulatory factor in tumour growth," Nat Rev Cancer **2**, 38–47 (2002).

[13] G. L. Semenza, "Targeting HIF-1 for cancer therapy," Nat Rev Cancer **3**, 721–732 (2003).

[14] G. L. Semenza, "Regulation of cancer cell metabolism by hypoxia-inducible factor 1," Sem Cancer Biol **19**, 12–16 (2009).

[15] G. L. Semenza, L. A. Shimoda, and N. R. Prabhakar, "Regulation of gene expression by HIF-1," (2006), Vol. 272, pp. 2–8.

[16] J. Kim and C. Dang, "Cancer's molecular sweet tooth and the Warburg effect," Cancer Res **66**, 8927–8930 (2006).

[17] A. M. Vannucchi and P. Guglielmelli, "Molecular pathophysiology of Philadelphia-negative myeloproliferative disorders: beyond JAK2 and MPL mutations," Haematologica **93**, 972–6 (2008).

[18] E. J. Baxter, L. M. Scott, and P. J. Campbell, "Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders," Lancet **365**, 1054–1061 (2005).

[19] P. J. Campbell and A. R. Green, "The myeloproliferative disorders," New Engl J Med **355**, 2452–2466 (2006).

[20] C. James, V. Ugo, J. P. Le Couédic, J. Staerk, F. Delhommeau, C. Lacout, L. Garon, H. Raslova, R. Berger, A. Bennaceur-Griscelli, J. L. Villeval, S. N. Constantinescu, N. Casadevall, and W. Vainchenker, "A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera," Nature **434**, 1144–1148 (2005).

[21] R. Kralovics, F. Passamonti, A. S. Buser, S. S. Teo, R. Tiedt, J. R. Passweg, A. Tichelli, M. Cazzola, and R. C. Skoda, "A gain-of-function mutation of JAK2 in myeloproliferative disorders," New Engl J Med **352**, 1779–1790 (2005).

[22] R. L. Levine, M. Wadleigh, J. Cools, B. L. Ebert, G. Wernig, B. J. P. Huntly, T. J. Boggon, I. Wlodarska, J. J. Clark, S. Moore, J. Adelsperger, S. Koo, J. C. Lee, S. Gabriel, T. Mercher, A. D'Andrea, S. Frhling, K. Dhner, P. Marynen, P. Vandenberghe, R. A. Mesa, A. Tefferi, J. D. Griffin, M. J. Eck, W. R. Sellers, M. Meyerson, T. R. Golub, S. J. Lee, and D. G. Gilliland, "Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis," Cancer Cell **7**, 387–397 (2005).

[23] R. Zhao, S. Xing, Z. Li, X. Fu, Q. Li, S. B. Krantz, and Z. J. Zhao, "Identification of an acquired JAK2 mutation in polycythemia vera," J Biol Chem **280**, 22 788–22 792 (2005).

[24] G. Chen and J. T. Prchal, "Polycythemia vera and its molecular basis: An update," Best Pract Res Clin Haematol **19**, 387–397 (2006).

[25] F. Delhommeau, S. Dupont, V. Della Valle, C. James, S. Trannoy, A. Massé, O. Kosmider, J. P. Le Couedic, F. Robert, A. Alberdi, Y. Lécluse, I. Plo, F. J. Dreyfus, C. Marzac, N. Casadevall, C. Lacombe, S. P. Romana, P. Dessen, J. Soulier, F. Viguié, M. Fontenay, W. Vainchenker, and O. A. Bernard, "Mutation in TET2 in myeloid cancers," New Engl J Med **360**, 2289–2301 (2009).

[26] F. H. Grand, C. E. Hidalgo-Curtis, T. Ernst, K. Zoi, C. Zoi, C. McGuire, S. Kreil, A. Jones, J. Score, G. Metzgeroth, D. Oscier, A. Hall, C. Brandts, H. Serve, A. Reiter, A. J. Chase, and N. C. P. Cross, "Frequent CBL mutations associated with 11q acquired uniparental disomy in myeloproliferative neoplasms," Blood **113**, 6182–6192 (2009).

[27] S. Swierczek, D. Yoon, C. Bellanne-Chantelot, S. Kim, C. Saint-Martin, F. Delhommeau, A. Najman, and J. Prchal, "Extent of hematopoietic involvement by TET2 mutations in JAK2V617F polycythemia vera," Haematologica **96**, 775–778 (2011).

[28] A. Tefferi and J. W. Vardiman, "Classification and diagnosis of myeloproliferative neoplasms: The 2008 World Health Organization criteria and point-of-care diagnostic algorithms," Leukemia **22**, 14–22 (2008).

[29] J. V. Melo and J. M. Goldman, *Hematologic Malignancies Myeloproliferative Disorders* (Springer, Berlin, 2007).

[30] G. Patrinos, B. Giardine, C. Riemer, W. Miller, D. Chui, N. Anagnou, H. Wajcman, and R. Hardison, *HbVar*, Improvements in the HbVar database of human hemoglobin variants and thalassemia mutations for population and sequence variation studies, `http://globin.cse.psu.edu/hbvar/menu.html`, acessed July 24 2011.

[31] P. Kleinert, M. Schmid, K. Zurbriggen, O. Speer, M. Schmugge, B. Roschitzki, S. S. Durka, U. Leopold, T. Kuster, C. W. Heizmann, H. Frischknecht, and H. Troxler, "Mass spectrometry: A tool for enhanced detection of hemoglobin variants," Clin Chem **54**, 69–76 (2008).

[32] C. Ou and C. Rognerud, "Diagnosis of hemoglobinopathies: Electrophoresis vs. HPLC," Clin Chim Acta **313**, 187–194 (2001).

CHAPTER 3

METHODS

3.1 Experimental methods

3.1.1 Confocal micro-Raman spectroscopy

Confocal microscopy is a technique that was developed to acquire images and spectroscopic information with high spatial resolution along the lateral[1] and the axial[2] directions. The resolution is, in principle, only diffraction limited. The basic idea of confocal microscopy is that the illumination and the detection beam paths have the same focal plane. In general, this condition is achieved by 2 pinholes placed at equal distances from the sample. The main advantage of confocal microscopy is the improved axial resolution as compared to wide field microscopy. Figure 3.1 depicts the principle of the setup of a confocal microscope; the pinhole on the detection path selects only the light emitted from the focal plane. Although resolution can be improved by the confocal configuration, care must be taken when selecting the pinhole diameter. If it is too small, it can substantially reduce the amount of light that reaches the detector. The detection pinhole ultimately provides an improvement in the axial resolution, and therefore, in cases where high axial resolution is not required, one can increase the pinhole's diameter, thus collecting more photons and reducing the intrinsic noise.

The equipment used for confocal imaging and microspectroscopy was the WITec alpha300 microscopy system (WITec, GmbH, Ulm, Germany). It is a modular microscope system that has a photon counting photomultiplier tube (PMT) and a spectrometer coupled to a charge couple device (CCD) camera, which in turn, allows

[1]Parallel to focus plane.

[2]Perpendicular to focus plane.

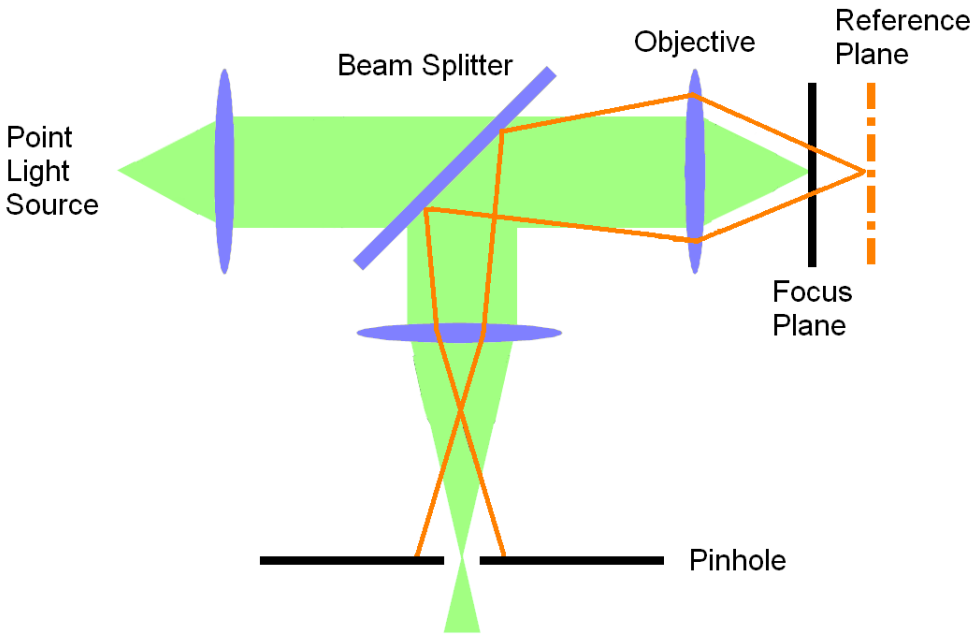


Figure 3.1: Principle setup for confocal microscopy. Modified from [1].

both confocal imaging and confocal raman microspectroscopy (See Figure 3.2). The illumination and detection pinholes are replaced by optical fibers that couple the laser to the upright microscope and the microscope to the spectrometer or the PMT. The equipment is also capable of scanning near field optical microscopy (SNOM) and atomic force microscopy (AFM). Neither of these will be discussed here.

In the actual configuration, light from a 488 nm Argon laser was focused onto the sample using a microscope objective. A 488 nm 'notch filter' was used to reject the elastic scattering, and unless otherwise specified, a 600 grooves/mm (500 nm blaze wavelength) diffraction grating was used to disperse the collected emission and obtain spectra in the 500 to 4500 cm^{-1} range using a CCD, operating at $-60^{o}C$ (Andor Technology, Belfast, Ireland). Unless otherwise specified, measurements were done under atmospheric conditions. Data were acquired and preprocessed using the WITec control software. Postprocessing was done with Matlab (The MathWorks, Inc. Natick, MA).

The microscope is also equipped with a piezoelectrically driven 3 axis translation

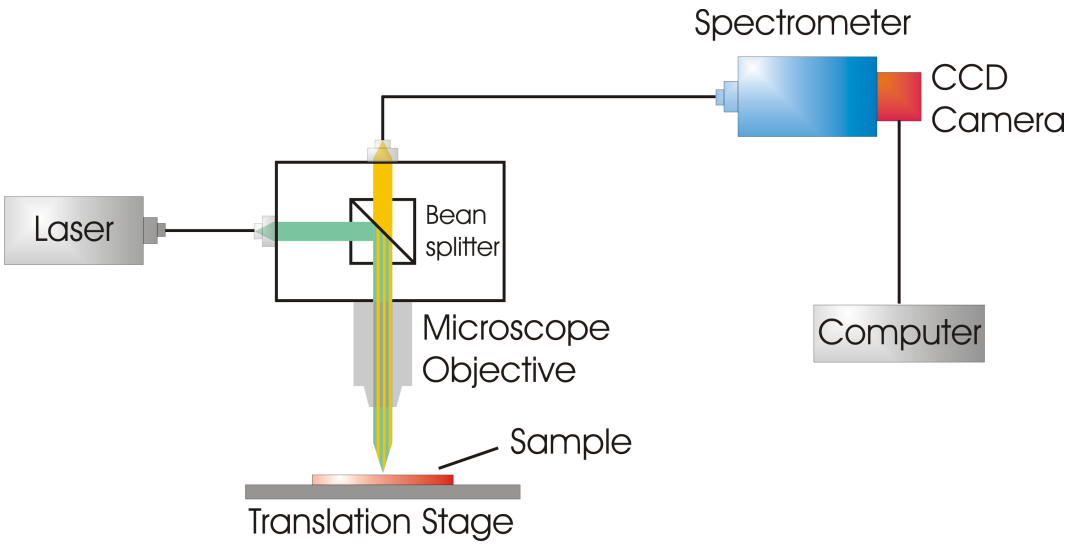


Figure 3.2: Experimental setup for confocal micro-Raman spectroscopy

stage that allows automatic scanning of the sample. The maximum scan range is $100\mu m$ parallel and $20\mu m$ perpendicular to the focal plane.

3.1.2 Silver nanoparticles substrates by Tollens reaction

The Tollens reaction is the deposition of semicontinuous silver films onto glass through the reduction of silver nitrate by an aldehyde. The first use of the Tollens reaction for the preparation of SERS active substrates was reported by Ni and Cotton in 1986 [2]. Since then, many of their properties have been elucidated. For example, it is known that silver aggregates produced by Tollens reaction organize in a fractal fashion making them attractive candidates for single molecule SERS [3].

One of the most appealing characteristics of substrates made by the Tollens reactions is the ease of preparation. The following protocol is based on published work by Wang *et al.* [3]. Stock solutions of 0.1 M silver nitrate ($AgNO_3$) and 0.8 M potassium hydroxide (KOH) in distilled water were prepared. A fresh solution of 0.5 M D-Glucose in distilled water was prepared for every batch. The $AgNO_3$, KOH, D-Glucose and ammonium hydroxide were analytical grade (Sigma-Aldrich, St. Louis, MO, USA).

For a final reaction volume V_f, $V_f/3$ of the 0.1 M $AgNO_3$ was prepared. While

stirring, the ammonium hydroxide was added one drop at a time to the $AgNO_3$ until a brown precipitate formed. The ammonium hydroxide was continually added until the solution was clear. The volume was adjusted to $V_f/3$. Then, $V_f/6$ of 0.8 KOH was added to the $AgNO_3$ solution forming a brown precipitate. Ammonium hydroxide was added dropwise until the solution became clear again. The glass substrates were soaked for 10 seconds in the former solution. Note: Glass was previously cleaned in organic solvents like ethanol and acetone to avoid carbon species contaminations. $V_f/2$ of the 0.5M glucose solution was mixed in the $AgNO_3$ solution prepared above. The glass stood in the final solution for 2 minutes. The size and separation of the silver nanoparticles will depend upon the time allowed for reaction [3]. The substrates were rinsed several times with warm and cold distilled water and blown dry using nitrogen or argon gas.

A scanning electron microscopy (SEM) image of a typical substrate with a reaction time of 2 minutes is shown in Figure 3.3A. An important feature of this type of substrates is self-similarity which has been characterized by the power law relationship between the area root mean square (RMS) of the silver nanoparticles and SEM imaging area of dimension $r \times r$ as shown by:

$$A_{RMS}(r) = kr^D \tag{3.1}$$

where D is the fractal dimension, and k is a constant related to the density of the fractal (i.e., Lacunarity) [4]. Figure 3.3B shows a loglog plot of A_{RMS} versus r, which demonstrates the self-similarity. Processing of the SEM images was done using the image processing software *ImageJ* [5]. Self-similarity is not a necessary nor a sufficient condition for electromagnetic enhancement although it often results in strong localization of surface plasmons. As a consequence of localization, self-similar substrates, such as Tollens, are capable of enhancement over a wide range of wavelengths [6].

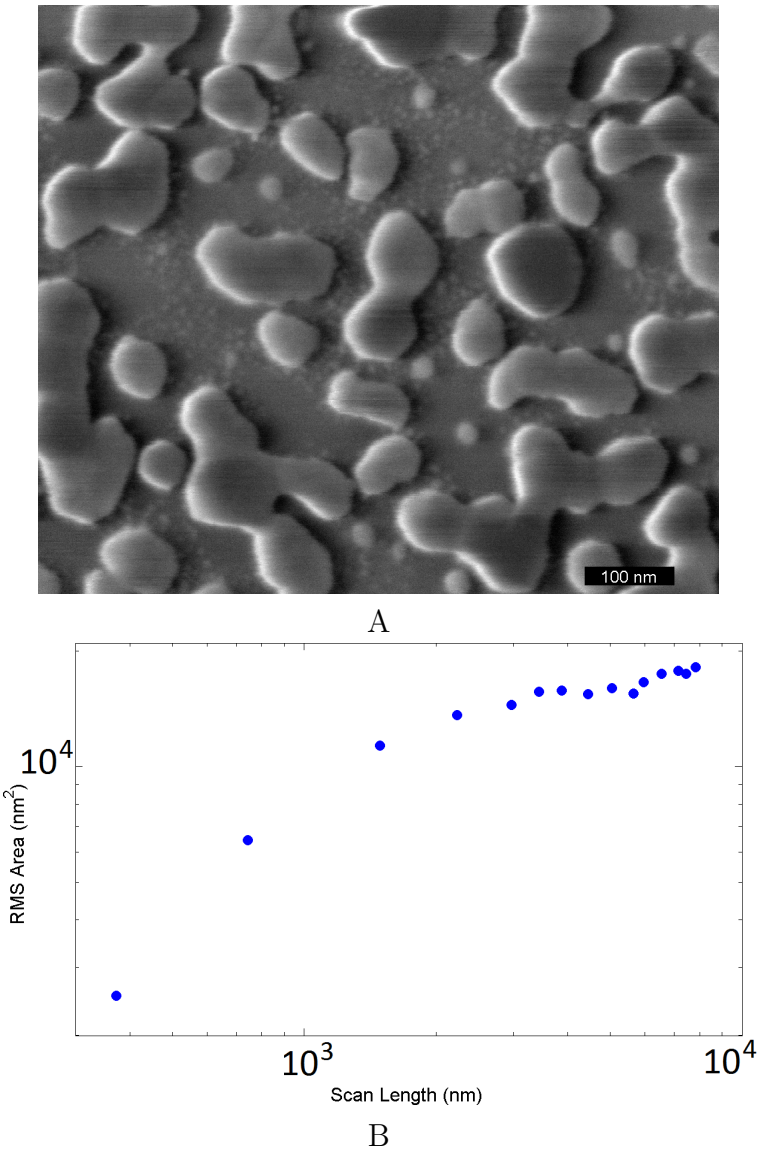


Figure 3.3: Tollens substrates characterization. **A:** SEM image of Tollens substrate with 120 seconds reaction. Scale bar is 100nm B **B:** Self-similarity behavior of the silver nanoparticles and aggregates result of the Tollens reaction.

3.1.3 Cell fraction separation by gradient media

Histopaque 1077 is a solution containing polysucrose and sodium diatrizoate, adjusted to a density of 1.077 g/mL that allows the isolation of granulocytes, platelets and mononuclear cell fractions by centrifugation. During the procedure, erythrocytes from anticoagulated blood are aggregated by polysucrose and rapidly sediment while granulocytes become slightly hypertonic, which increases their sedimentation rate, resulting in pelleting at the bottom of the centrifuge tube. Lymphocytes and other mononuclear cells remain at the plasma/Histopaque interface [7] (see Figure 3.4). The reagents and procedure used are as follows:

Reagents

- **10X Lysis Buffer:** The following is the recipe for a 10X stock solution of lysis buffer. This buffer has been formulated for erythrocytes lysis while having minimum effect on leukocytes. Mix all the ingredients listed below and adjust volume to 100ml with autoclaved distilled water. Filter solution. Store at room temperature.
  - 8.29grams NH_4CL.
  - 0.79grams NH_4HCO_3.
  - 2ml 0.5M Ethylenediaminetetraacetic acid (EDTA) pH 8.0.

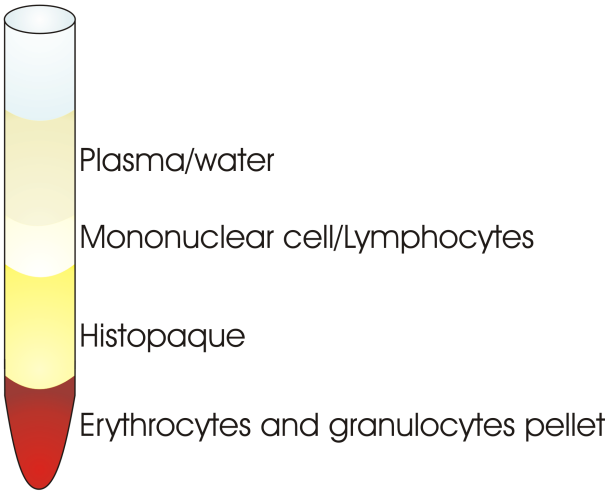


Figure 3.4: Layers form in tube after centrifugation of anticoagulated blood with histopaque 1077.

- Phosphate buffered saline (PBS).
- Histopaque 1077(Sigma-Aldrich, St. Louis, MO, USA).

Procedure

1. Draw 10mL of peripheral blood by venipuncture in acid citrate dextrose(ACD).
2. Centrifuge blood at 700 rpm for 10 min.
3. Remove upper plasma layer that contains platelets. It will be about 1 cm thick. Put it in a clean eppendorf tube.
4. **Platelets:**
   (a) Centrifuge plasma at 400 g for 10 min.
   (b) Pipet plasma out. If pellet (platelets) red, add 1ml of lysis buffer and let stand for 10 min.
   (c) Centrifuge at 400 g for 10 min. Discard supernatant.
   (d) Repeat lysis as necessary until pellet is white.
   (e) Discard supernatant and add 1ml of PBS. Pipet to mix.
   (f) Centrifuge at 400 g for 10 min.
   (g) Repeat wash. Discard supernatant.
   (h) Resuspend platelets to final concentration.
5. Dilute the lower thick red layer at least 1:1 with PBS. If blood is from PV patient, dilute even more.
6. Add 3 ml of sterile histopaque into each 15 ml centrifuge tube (one tube for each 8ml of diluted blood).
7. Very slowly add the diluted blood onto the histopaque, making sure they do not mix. Add no more than 8 ml of diluted blood.
8. Centrifuge at 400 g for 30 min. Ensure NO BRAKE on centrifuge.
9. **Mononuclear cells:** After centrifugation at 400 g remove the cloudy layer that contains the mononuclear cells. Put in a clean 15 ml centrifuge tube (no more than two layers per tube) and fill tube to 15ml with PBS. Mix well.

(a) Spin down at 400g for 10 min.

(b) Pipet out supernatant. Discard.

(c) Add 2 ml of lysis buffer. Let stand for 10 min.

(d) Centrifuge at 400 g for 10 min. If pellet is red. Repeat lysis.

(e) Discard supernatant. Fill tube with PBS and mix well.

(f) Centrifuge at 400 g for 10 min.

(g) Repeat wash.

(h) Resuspend cells in PBS to final concentration.

10. **Granulocytes:** Remove the remaining volume to the cell pellet which should be red and occupy about 2ml.

(a) Add 4 volumes of lysis buffer. Pipet to mix. Let stand for 10 min at room temperature.

(b) Centrifuge at 2500 rpm for 10 min.

(c) Decant supernatant.

(d) Repeat lysis as necessary (usually 2 or 3 times) until pellet is white,centrifuging at 450 g for 10 min each time.

(e) Decant supernatant.

(f) Fill tube to 15 ml with PBS and mix well.

(g) Centrifuge at 450 g for 10 min.

(h) Repeat wash.

(i) Resuspend pellet in PBS to final concentration.

3.1.4 Reverse transcription real time polymerase chain reaction (qRT-PCR)

Real time polymerase chain reaction (PCR), in conjunction with reverse transcription, is a very effective tool to quantify gene expression from ribonucleic acid (RNA). PCR is a process for (DNA) amplification able to generate large amounts of the DNA sequence in question ($\sim 10^5$copies [8]). The overall process can be divided into three steps: 1) denaturation, 2) annealing and 3) elongation. During the first step, double stranded DNA is heated to 94^oC. At this temperature the strands will

separate (denature) allowing *primers*[3] to access the single stranded DNA (ssDNA). The annealing step consists of cooling the reaction mixture to about 50^o. This allows *primers* to select and bind to the single stranded DNA template molecules. Finally, during the elongation step, the reaction solution is heated to 72^o, and in the presence of all *nucleotide bases*, the heat stable *polymerase* will drive the replication by reading the template DNA and synthesizing the new single stranded DNA. Each repetition of the process will double the amount of DNA.

Amplification of messenger RNA (mRNA) is necessary when quantification of the expression of a gene is desired. However, RNA is not suitable for amplification by the DNA polymerases used for PCR, thus making reverse transcription indispensable. RNA can be reverse transcribed into its complementary DNA (cDNA) using *reverse transcriptase*, an enzyme usually obtained from retroviruses.

Unfortunately, accurate quantification of DNA is not possible with normal PCR, as the final amount of DNA does not necessarily mirror the initial amount of DNA present in the reaction. This phenomenon is due to many factors including, but not limited to, deactivation of the polymerase, shortage of nucleotide bases and/or primers, reannealing of amplified DNA, etc. [9]. To overcome this issue, real time PCR also known as quantitative PCR (qPCR) was developed in the 90s [10]. The principle of qPCR is the introduction of a fluorescent probe with two fluorophores (reporter and quencher) that upon polymerase cleavage, increases fluorescence. See Figure 3.5.

The fluorescent signal allows real time detection of the amplification process, and a sigmoidal curve of fluorescence versus amplification cycles is produced see Figure 3.6A. During the first cycles no increased in the fluorescence is detectable. This part of the graph is called the baseline and is used as the background of the measurement. Thus, subtraction of the baseline from the measurements yields the absolute increased in fluorescence. The second part is an exponential increase of the fluorescence. Finally the curve goes into a nonexponential plateau at higher cycle numbers that occur due to

[3]ssDNA, usually 18 to 28 nucleotides in length [9] that serve as the starting point for DNA replication.

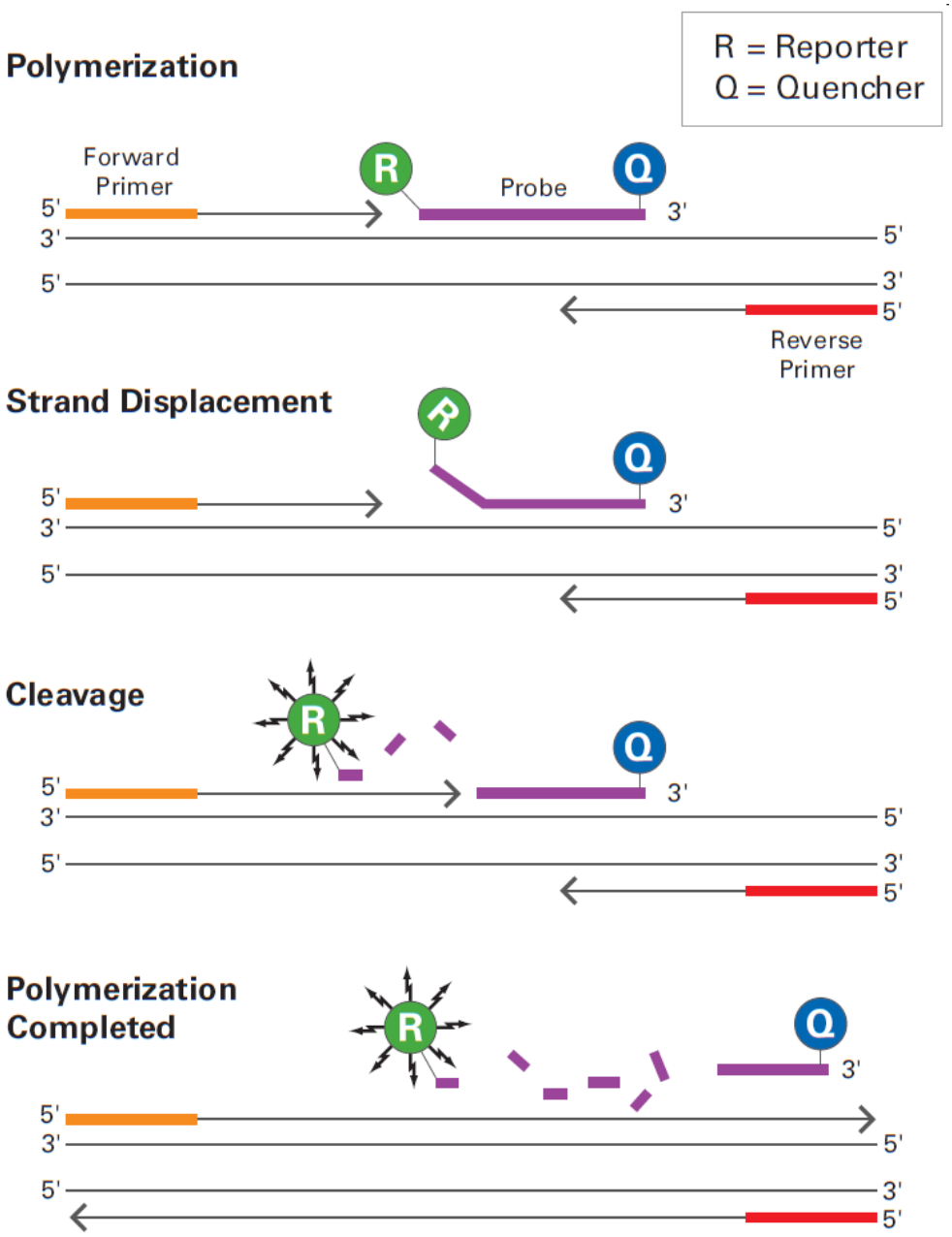
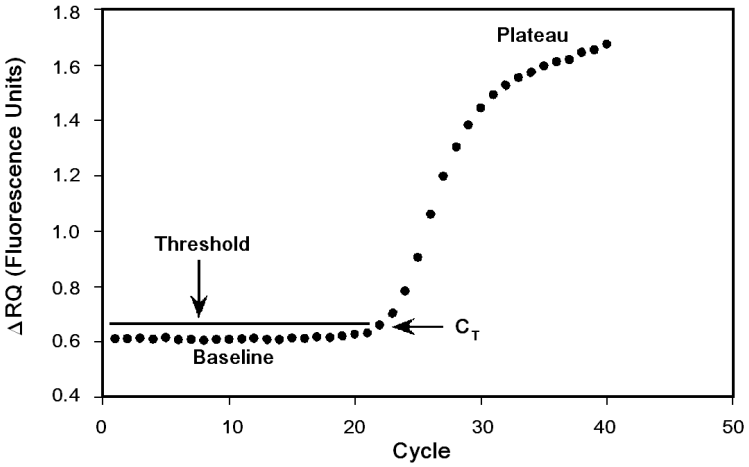
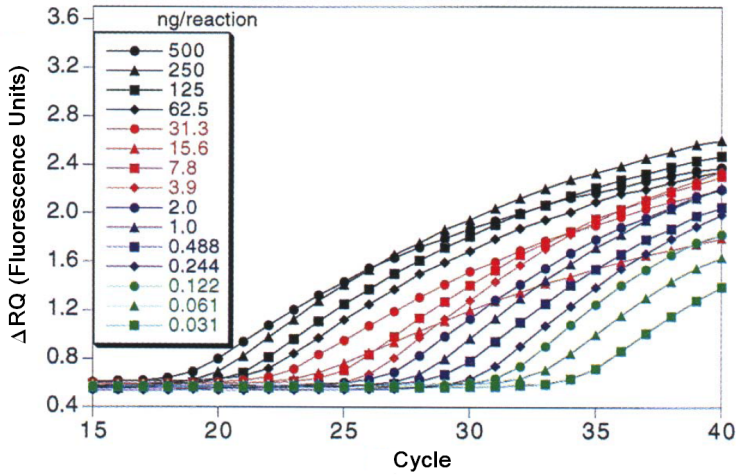


Figure 3.5: Mechanism of the qPCR probe cleavage that results in a fluorescent signal. Reproduced with permission from [11]. Copyright 2010, Life Technologies Corporation.



A



B

Figure 3.6: Amplification plots of **A:** Real time quantitative PCR; and **B:** 1:2 serially diluted samples of human DNA samples. Reprinted with permission from [10]. Copyright 1996, Cold spring harbor laboratory press.

the exhaustion of reaction components (primers, probes, nucleotide bases, etc.). The amplification plot provides information that allows determination of the quality of the measurement being done and the final result, which is given as a cycle number (C_T) at which the enough amplified material accumulates to yield a detectable increased in the fluorescence (see Figure 3.6A). C_T value is proportional to the initial quantity of target DNA used for the amplification [10], see Figure 3.6B.

C_T values mainly allow for two types of quantification, absolute and relative. Absolute quantification is used to determine the number of copies of a particular gene present in a sample. This requires normalization to an initial known quantity such as the number of cells or the amount of nucleic acid. Relative quantification, on the other hand, determines the target gene expression relative to a reference gene.[4] The advantage of this method is that it does not require quantification of the initial amount of material. The pitfall is the need for a reference gene that must have a constant expression among all the samples tested and for which expression must not be affected by the different conditions tested in the study such as treatment or the presence of a disease.

Relative quantification is done by measuring C_T for both the target and the reference genes in the same sample. This allows compensation for differences in the initial amount of material for each sample. Additional normalization can be done with respect to a calibrator sample or set, such as healthy (control) or untreated samples. Therefore, the final result will be given in terms of a fold change in normalized gene expression of one group (i.e., patients) with respect to another (i.e., controls).

Gene expression of glucose transporter 1 (SLC2A1), vascular endothelial growth factor A (VEGF-A) and pyruvate dehydrogenase kinase, isozyme 1 (PDK1) relative to the hypoxanthine phosphoribosyltransferase 1 (HPRT1) housekeeping gene were measured from RNA from blood cells. The selection of HPRT1 as housekeeping gene is based on its reported constant expression in neutrophils [12], reticulocytes [13] and leukocytes, in general [14].

---

[4]Also known as housekeeping gene.

Total RNA from granulocytes was isolated by Tri-Reagent (Molecular Research Center, Cincinnati, OH). Total RNA (50 ng) was reverse transcribed using SuperScript III First-Strand Synthesis SuperMix for qRT-PCR (Invitrogen, Carlsbad, CA). RNA degradation was checked by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). RNA with well defined 28s and 18s bands in SDS-PAGE was used. Additionally, the quantity and quality of RNA were checked using spectrophotometric analysis (NanoDrop, Thermo Fisher Scientific Inc., DE, USA) by measuring absorption at 230 nm, 260 nm and 280 nm.

Expression of the genes was determined by qRT-PCR using TaqMan Expression Assays (SLC2A1: Hs00892681_m1; VEGF-A: Hs99999070_m1; PDK1: Hs00176853_m1; HPRT1: 4333768F) (Applied Biosystems, Foster City, CA). Relative gene expression was evaluated using Biogazelle QBasePLUS (Zwijnaarde, Belgium).

3.1.5 Samples

"It is the policy of the Institutional Review Board (IRB) that all human subjects research that is conducted by or under the direction of any employee, faculty, staff, student, or agent of the University of Utah in connection with his or her institutional responsibilities must be reviewed by the IRB."[5] The protocols for the research presented as part of the present doctoral thesis have been reviewed and approved by the University of Utah IRB.

3.1.5.1 Human whole blood samples from myeloproliferative patients

This study was previously reviewed and approved by The University of Utah IRB as a minimal risk study on 6/10/2009 (IRB No. 35136, PI: Josef Prchal). Fresh blood samples from consented patients with known myeloproliferative disorders (Essential Thrombocythemia, Polycythemia Vera and Idiopathic Myelofibrosis) were collected at the University of Utah Hospital and VA Salt Lake City Medical Center. Protocol was revised to include control samples and reapproved on May 2010. A second revision of

[5]University of Utah Institutional Review Board Website.

the protocol was done to include secondary polycythemia and thrombocytosis patients who were screened as part of IRB No. 17793 and reapproved on May 2011 .

3.1.5.2 Human serum and peripheral mononuclear cells before and after influenza vaccination

This study was previously reviewed and approved by The Tniversity of Utah IRB as nonhuman subject research on 10/16/2009 (IRB No. 38405, PI: Valy Vardeny) to obtain deidentified serum and peripheral blood mononuclear cells (PBMCs) from patients with heart failure and controls administered during the 2009/2010 influenza season banked through the IRB project "The use of high dose influenza vaccine compared to standard dose vaccine in patients with heart failure - a pilot study" submitted to the University of Wisconsin.

3.1.5.3 Whole blood for hemoglobin variants

Deidentified samples of whole blood from patients affected with hemoglobinopathies were obtained from ARUP Laboratories (Salt Lake City, Utah) under the approved IRB protocol "Blood, urine and other body fluid samples for validation of assay methods used in clinical testing" (IRB No. 7275).

3.2 Statistical methods

3.2.1 Hypothesis testing

Hypothesis testing is a statistical tool used to draw inferences about populations from sets of empirical data. Hypothesis testing can be done on one or multiple sets of samples. One-sample tests are used to test if a population's parameter, usually the mean or the median, is different from a set value. On the other hand, two or more sample sets tests are used to test whether the compared populations have different parameter values. Hypothesis testing starts by making a statement about the populations that are to be tested. This statement is called the *null hypothesis*(H_0), and it is, in principle, an assertion that one can only disprove. That means that the test has only one of two possible outcomes: rejecting H_0 or failing to reject H_0. A parameter of a population can be tested under the hypothesis that it is less or larger than some value. Such a test is known as a *one-tailed test*, as the region of

rejection is only one side of the sampling distribution. On the other hand, when the hypothesis is that the parameter is exactly equal to some value, the test is said to be *two-tailed*. Interpreting results from hypothesis testing is often done in terms of the *p-value*. P-value is the probability that, for other sampling, one obtains data at least as extreme as the one that was observed, assuming that the null hypothesis is true. Therefore, small p-values suggest H_0 is unlikely to be true. H_0 is rejected if the p-value is less than the significance level (i.e., critical p-value). Among the scientific community, a 5% or 1% significance level is often adopted.

There are multiple methods for hypothesis testing. They differ in the assumption made *a priori* about the distributions of the variables under study. Most common tests assume normal distributions and are called parametric tests. Another common assumption is homoscedasticity (equal variances of the tested distributions). Before using any of the tests, it should be clear whether the data complies with the assumption made by the them. This can be done by testing normality and homoscedasticity using the Kolmogorov-Smirnov test and the Levene's tests, respectively. Table 3.1 shows some of the tests and their respective assumptions.

Hypothesis testing is used in Chapter 4 to test whether spectroscopic variables measured on healthy and diseased samples come from the same population, meaning the *null hypothesis is that healthy and diseased populations have the same mean or median* and that we should used two-tailed tests. Variables were tested for normality and homoscedasticity. Either Kruskal-Wallis or the t-test for unequal variances was used, accordingly. We repeat this procedure for thousands of variables measured on

Table 3.1: Hypothesis testing methods and their assumptions made on the distributions of the variables tested.

Test	Parameter Tested	Assumptions
Analysis of variance	Means	Normality and homoscedasticity
t-test	Means	Normality and unknown but equal variances
t-test(unequal variances)	Means	Normality
Kruskal-Wallis	Medians	Equal continuous distributions

the same set of samples. This introduces an additional problem described as follows: The more tests we performed on a data set, the more likely we were to reject the null hypothesis when it is true [15]. To overcome this problem one must adjust the significance level and make it more stringent. One way to do this is using the *Šidk-Bonferonni* approximation for nonindependent tests, and the new significant level becomes:

$$\alpha_{adj} = \frac{\alpha}{C} \tag{3.2}$$

where C is the number of tests done on the set of data.

3.2.2 Receiver operating characteristic curve

Receiver operating characteristic (ROC) curves are a beautiful way to visualize binomial (positive or negative) classifiers based on their performance, which is often defined in terms of sensitivity and specificity. Sensitivity is a measure of the classifier's ability to identify positive cases (eq. 3.3), while specificity is a measure of the ability to identify negative cases (eq. 3.4):

$$sensitivity = \frac{True\ Positives}{True\ Positives + False\ Negatives} \tag{3.3}$$

$$specificity = \frac{True\ Negatives}{True\ Negatives + False\ Positives} \tag{3.4}$$

The ROC curve is a plot of sensitivity (i.e., True positive rate) versus 1-specificity (i.e., False positive rate) for different cut-off points, and thus it renders the tradeoff between sensitivity and specificity. The point (0,1) represents the ideal classifier. That is, no false negatives or false positives are assigned by the classifier. Since in practice this is not achieved, the best cut-off point is defined as the point on the ROC curve with the shortest distance away from the ideal classifier. Figure 3.7 shows a sample ROC curve where the diagonal line represents a random classifier and the blue point is the best cut-off.

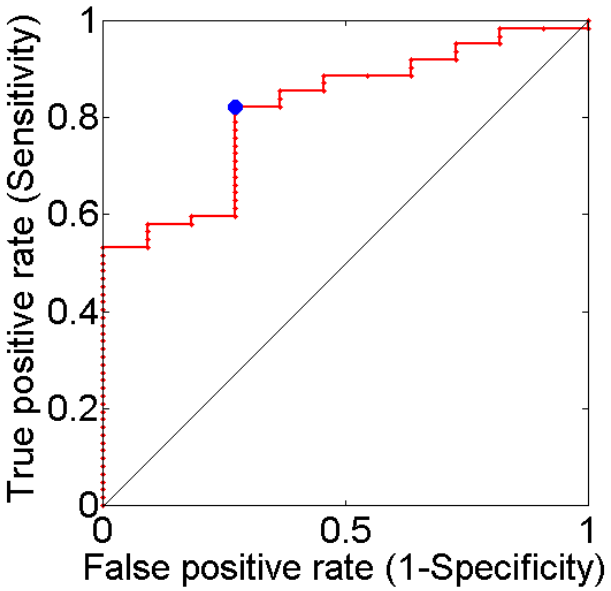


Figure 3.7: Example of an ROC curve (red). Diagonal line corresponds to random classifier. Blue dot is the cut-off point defined as the closest to the ideal test point.

In order to compare the performance of multiple classifiers or tests, the area under the curve (AUC) is calculated. The AUC has values between 0 and 1, but since the random classifier (defined by the diagonal of the unit square) has an AUC=0.5, a useful test should never have an AUC lower than 0.5.

The AUC is equivalent to the probability that the test will rank a randomly chosen positive case higher than a randomly chosen negative case [16]. Thus, in general, the higher the AUC is the better the classifier's performance is. Table 3.2 shows the standard adopted by the medical sciences community to evaluate a classifier's performance based on AUC values [17–19].

In Chapter 4, ROC curves are use to evaluate the performance of various spectral variables to distinguish among healthy and diseased subjects. Unless otherwise noted, diseased subjects are defined as positive and healthy subjects are defined as negative cases.

Table 3.2: ROC area under the curve based performance evaluation.

AUC	Classifier Performance
0.9-1	Excellent
0.8-0.9	Good
0.7-0.8	Fair
0.6-.07	Poor
<0.6	Failed

3.2.3 Multivariate analysis

Spectroscopic data are multivariate in nature, meaning that the underling phenomenon is described by more than one variable. In the spectroscopic case, we have values (I) measured in units of light intensity or photon count, for a set of n wavelength, frequency or energy points. If each $I(\omega_n)$ is considered a variable, a spectrum is then a multivariate set and multiple spectra can be analyzed under the multivariate analysis (MVA) framework.

Multivariate data analysis can be classified in three groups according to its final purpose [20]: Data description, discrimination and classification, or regression and prediction. **Data description** is concerned with the exploration and characterization of the data that could be done in term of means and standard deviations or correlations between variables, as well as finding which variables are the most relevant to describe the phenomenon under scope. **Discrimination and classification** of data are done to separate and to classify samples based on a set of variables and some *a priori* knowledge of the samples (e.g., healthy or diseased sample). The most popular technique for data description and classification is principal component analysis, and it is explained in more detailed below. Finally, a mathematical model or regression can be constructed based on variables measured on a training set of samples from which some additional knowledge is available. It can then be applied to a new set of samples to calculate an indirect property. Some regression and prediction techniques are Principal Components Regression and Partial Least Squares Regression.

3.2.3.1 Principal components analysis

Principal component analysis (PCA) is a very powerful application of linear algebra. One of its more important strengths is that it is nonparametric. As such, PCA results and capabilities do not change upon variables' distributions. The main objective of PCA is to compute the most relevant algebraic basis to represent and to attain the most information out of the data.

PCA works as a linear transformation from the space of variables with m dimensions (where m is the number of variables, and each sample is represented as an $m \times 1$ vector) to a new space that represents the data in a more meaningful way. The problem is that there are many changes of basis that one could use. This problem is solved by assuming that the directions within the data with maximum variance are the ones that carry most of the information. Figure 3.8A shows a set of data represented in a 3-dimensional space of variables. The data set displays most of the variability along the line depicted in red. Such a line would constitute the first axis of the new space, also known as the first principal component.

Mathematically, PCA starts with the construction of the matrices $X_{m \times n}$, where each column corresponds to the set of measurements of a sample and the covariance matrix C_X (eq. 3.5) [21]. C_X is a square symmetric $m \times m$ matrix whose diagonal terms correspond to the variances of the m variables and the off-diagonal elements to the covariances between different variables.

$$C_X = \frac{1}{n-1} X X^T \tag{3.5}$$

The optimal basis is the one that makes C_X diagonal, that way all newly defined variables will be independent from each other, and variances of each new variable will be maximized. In the basis where C_X is diagonal, the basis vectors correspond to the principal components and they will lie in the directions of maximum variance, starting with PCA_1. So the objective is to find a matrix $\mathbf{P}$, where $\mathbf{Y} = \mathbf{PX}$, such that $C_Y = \frac{1}{n-1} Y Y^T$ is diagonal. The rows of $\mathbf{P}$ are the principal components (PCs) of $\mathbf{X}$ [21]. It is useful to subtract the mean of each variable from each row of $X_{m \times n}$,

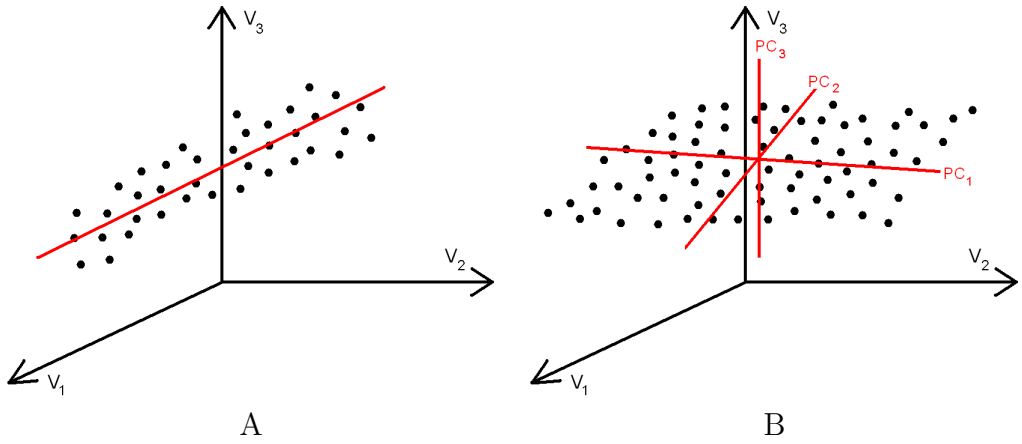
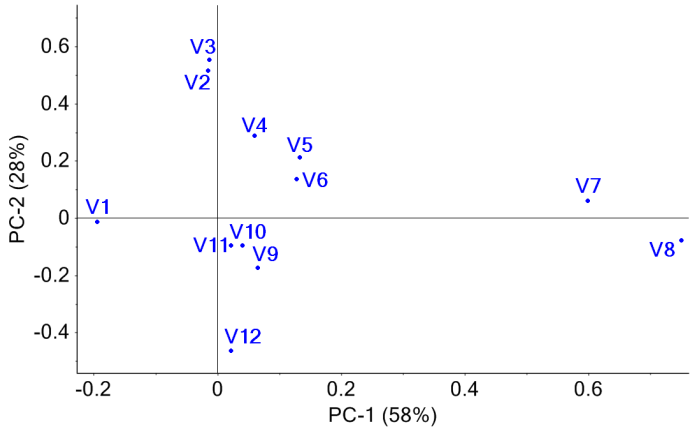


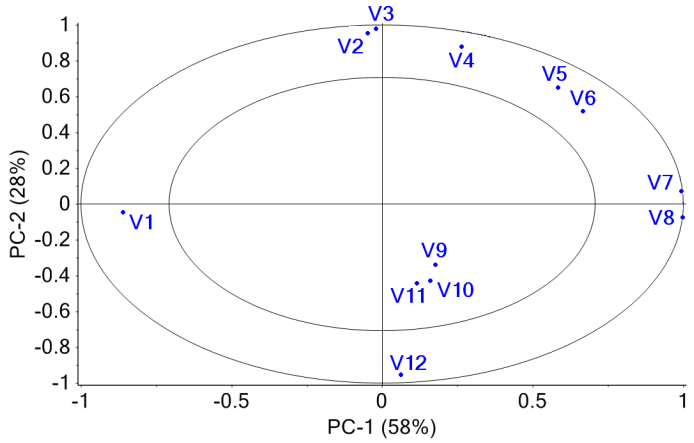
Figure 3.8: Principal component analysis. **A:** Data points in a 3-dimensional variable space. Red line shows the direction of most variability within the data. **B:** Different set of data point that show a different distribution in the variables space. Whereas in A most of the variance can be describe along one direction (i.e., one PC), in B two PCs are required.

making the origin of the PCs the center of gravity of the data as depicted in Figure 3.8B in the case of 3-dimensional space. One of the most useful features of PCA is the reduction of dimensionality of a data set, that is, to elucidate the variables that mostly contribute to describe the variance among the data. For example, in Figure 3.8B, it is observable that most of the variance in the data is along the first and second principal components, while the third is almost negligible. In that case we have been able to go from a 3-dimensional to a 2-dimensional space, which might seem not that useful. However, in Chapter 5 we will show that PCA analysis allows reduction from a 1024-dimension space to a 2-dimensional space.

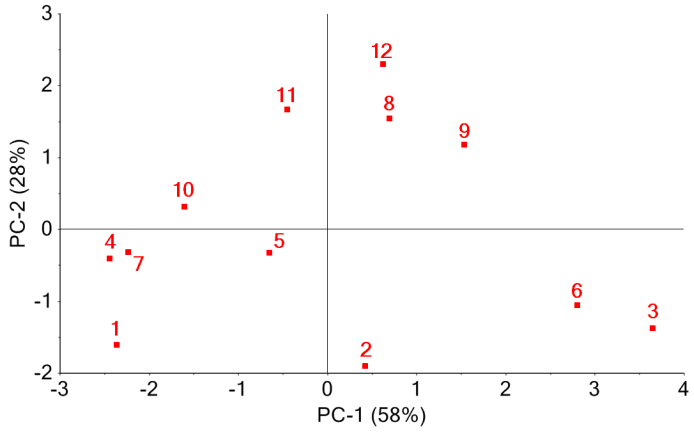
In practice PCA is done with help of two very useful plots: the loadings and the scores plot. The PCs are vectors in variable space that are expressed as a linear combination of the unit vector (i.e., each variable vector). The coefficients that describe that linear combination are the loadings such that each PC will have m corresponding loadings. The loading plot is usually a 2-dimensional map of the variables in which each variable is represented by a pair of loadings (see Figure 3.9A). The loading plot provides visualization of the intervariable relationships [20],



A



B



C

Figure 3.9: Examples of PCA plots used for the analysis of data. **A:** Loadings or variables map for two PCs. **B:** Scale independent variables plot used to assess intervariable relations. **C:** Scores or samples plot for the same PCs as in A. PCA was done using *The Unscrambler* [22] with software's example data set.

but a more useful plot is the correlation loadings plot which depicts loadings in a scale-independent fashion so relations between variables are easily assessed. Figure 3.9B is the correlation loadings plot of the data in Figure 3.9A. The variable V8 is the one that contributes the most to PC1 whereas V3 is the one that most contributes to PC2, variables V9, V10 and V11 contribute very little to either PC. Additionally, one observes from 3.9B that V7 and V8 are highly positively correlated while V3 and V12 are highly negatively correlated. From 3.9A and 3.9B one could conclude that variables V2, V3, V7, V8 and V12 will suffice to describe most of the variability among the data.

A scores plot or samples mapping depicts each sample in the PCs space. Objects close to the origin are the most typical whereas the ones far from it are extremes. Objects close to each other in the scores plot are similar. This property of the scores plot makes PCA viable for discrimination and classification of data. PCA is used in Chapter 5 for the analysis and classification of multiple SERS spectra.

3.3 Acknowledgements

We thank Dr. Matt Delong for assistance with SEM images.

3.4 References

[1] WITec Wissenschaftliche Instrumente und Technologie GmbH, Horvelsinger Weg 6, 89081 Ulm Germany, *Confocal Microscopy Configuration Manual* (2007).

[2] F. Ni and T. M. Cotton, "Chemical procedure for preparing surface-enhanced Raman scattering active silver films," Anal Chem **58**, 3159–3163 (1986).

[3] Z. Wang, S. Pan, T. D. Krauss, H. Du, and L. J. Rothberg, "The structural basis for giant enhancement enabling single-molecule Raman scattering," P Natl Acad Sci USA **100**, 8638–8643 (2003).

[4] O. Malcai, D. A. Lidar, O. Biham, and D. Avnir, "Scaling range and cutoffs in empirical fractals," Phys Rev E **56**, 2817–2828 (1997).

[5] W. Rasband, *ImageJ*, National Institutes of Health (NIH), v1.36b edn.

[6] C. Douketis, T. L. Haslett, Z. Wang, M. Moskovits, and S. Iannotta, "Self-affine silver films and surface-enhanced Raman scattering: Linking spectroscopy to morphology," J Chem Phys **113**, 11 315–11 323 (2000).

[7] Sigma-Aldrich, *Product Information Histopaque 1077*, St Louis, MO, USA.

[8] J. Walker and R. Rapley, *Medical Biomethods Handbook* (Humana Press, Totowa, NJ, 2005).

[9] G. Viljoen, L. H. Nel, and J. R. Crowther, *Molecular Diagnostic PCR Handbook* (Springer, Dordrecht, The Netherlands, 2005).

[10] C. A. Heid, J. Stevens, K. J. Livak, and P. M. Williams, "Real time quantitative PCR," Genome Res **6**, 986–994 (1996).

[11] A. Biosystems, *TaqMan Gene Expression Assays Product Overview*, Life Technologies Corporation, Foster City, CA, USA, taqMan is a registered trademark of Roche Molecular Systems, Inc.

[12] X. Zhang, L. Ding, and A. J. Sandford, "Selection of reference genes for gene expression studies in human neutrophils by real-time PCR," BMC Mol Bio **6** (2005).

[13] N. Silver, S. Best, J. Jiang, and S. L. Thein, "Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR," BMC Mol Bio **7** (2006).

[14] G. Spinsanti, R. Zannolli, C. Panti, I. Ceccarelli, L. Marsili, V. Bachiocco, F. Frati, and A. Aloisi, "Quantitative real-time PCR detection of TRPV1-4 gene expression in human leukocytes from healthy and hyposensitive subjects," Mol Pain **4** (2008).

[15] N. Salkind, *Encyclopedia of Measurement and Statistics* (SAGE Publications, 2007).

[16] T. Fawcett, "An introduction to ROC analysis," Pattern Recogn Lett **27**, 861–874 (2006).

[17] P. Brubaker, "Do not be statistically cenophobic: TIME TO ROC AND ROLL!!" J Cardiopulm Rehabil **28**, 420–421 (2008).

[18] S. Hansson, A. Svanstrom-Rojvall, M. Rastam, C. Gillberg, C. Gillberg, and H. Anckarsater, "Psychiatric telephone interview with parents for screening of childhood autism - tics, attention-deficit hyperactivity disorder and other comorbidities (A-TAC): Preliminary reliability and validity," Brit J Psychiat **187**, 262–267 (2005).

[19] M. Osuchowski, K. Welch, J. Siddiqui, and D. Remick, "Circulating cytokine/inhibitor profiles reshape the understanding of the SIRS/CARS continuum in sepsis and predict mortality," J Immunol **177**, 1967–1974 (2006).

[20] K. Esbensen, D. Guyot, F. Westad, and L. Houmller, *Multivariate Data Analysis In Practice : An Introduction to Multivariate Data Analysis and Experimental Design* (Camo, Oslo, 2009).

[21] J. Shlens, *A Tutorial on Principal Component Analysis*, Systems Neurobiology Laboratory, Salk Insitute for Biological Studies, http://www.snl.salk.edu/ shlens/pca.pdf (2005).

[22] C. S. AS, *The Unscrambler X* (Oslo, Norway, 2009-2010), build version: 0.0.0.42 edn.

CHAPTER 4

SURFACE ENHANCED RAMAN SPECTROSCOPY ON MYELOPROLIFERATIVE NEOPLASMS[1,2]

4.1 Introduction

Myeloproliferative neoplasms (MPNs) belong to a family of diseases that originate in a somatic mutation in a self-renewing hematopoietic stem cell. MPNs include four main diseases: chronic myelogenous leukemia (CML), which is characterized by Philadelphia chromosomal abnormality; polycythemia vera (PV); essential thrombocythemia (ET); and primary myelofibrosis (PMF). In this study, we focused on so called Philadelphia chromosome negative disorders: PV, ET and PMF. Characteristics of these diseases are bone marrow hypercellularity, predisposition to thrombosis and hemorrhage, and an increased risk of evolution towards universally fatal acute leukemia [1, 2]. These three MPNs share, in a variable proportion of each group, a somatic clonal mutation of a tyrosine kinase gene that encodes for *JAK2* V617F [3]. While this mutation is diagnostic for the presence of an MPN, but not its type, its absence does not exclude the existence of MPN [4]. In JAK2 V617F negative patients, the diagnosis of PV and ET may be challenging and in some instances, may be difficult and expensive to differentiate from reactive polycythemic, nonmalignant disorders and congenital polycythemias, whereas ET is largely diagnosed by exclusion of other conditions with increased platelet count due to the 'lack of specific tests' [4, 5]. More

---

[1] Parts of this chapter were submitted in September 2011 as an article to *Science*, coauthored with Professor Valy Vardeny, Professor Sabina Swierczek and Doctor Josef T. Prchal from the Division of Hematology, School of Medicine at the University of Utah.

[2] Parts of this chapter resulted in US patent application 61/428,697 (December 2010, see Appendix) coauthored with Doctor Josef T. Prchal and Professor Valy Vardeny.

detailed molecular basis of MPNs and current diagnosis strategies are both found in Section 2.4.

In recent years, Raman spectroscopy has proven to be a successful diagnosis technique of a wide range of diseases including atherosclerosis, kidney stones, bone diseases, diabetes, and a broad collection of neoplasms [6]. Optical spectroscopy has several advantages over more traditional diagnosis methods (i.e., histopathology, quantitative PCR, etc.) such as faster data analysis; nonspecific sample preparation; nonspecific labels/reagents/antibodies requirements; and immediate on-site implementation.

Raman scattering can be described as the inelastic scattering of an impinging photon by a molecule. The energy shift of the scattered photon usually corresponds to a vibrational mode of the molecule. Various Raman active vibrational modes cause different photon energy shifts. Thus, the Raman scattering spectrum is highly specific to a molecule, as each type of molecule has a distinctive set of vibrational modes (see Section 1.2). However, the Raman scattering cross section is very small, and this usually requires an enhancement mechanism. Raman scattering can be enhanced in the vicinity of subwavelength size metallic nanoparticles or rough surfaces [7], giving rise to surface enhanced Raman scattering (SERS). The enhancement is mainly electrostatic in nature, although there is a chemical contribution that results from the adsorption of molecules onto the metal surface [8](see Section 1.3).

As describe in section 2.3, the so called Warburg effect is associated with increased activity of hypoxia regulated transcriptional factors, i.e., HIF-1 and HIF-2 that bind to evolutionarily conserved oligonucleotides of their target genes. It is an adaptive response that leads to augmentation of aerobic glycolysis in cancer tissues, which in turn, makes cancer cells consume more glucose under aerobic conditions than normal cells. This is a response mediated by the hypoxia induced transcription factor (HIF) [9].

The analysis of SERS spectra from granulocytes reveals that it is possible to detect differences among MPN patients and healthy controls. This represents a *radical change* in the diagnostic approach currently used for differentiation of neoplastic

from reactive and congenital disorders. The experimental conditions for the SERS measurements and SERS spectrum analysis can be easily achieved, laser induced damage to the sample can be neglected, and results may be rapidly obtained. In this chapter, we show that SERS background continuum carries very valuable information that allows optical diagnosis of MPN. The described technique involves straight forward substrate and sample preparation, and allows fast diagnosis. Moreover, we provide supporting evidence that MPN, similar to other malignancies, modify their metabolism by the Warburg effect.

4.2 Study subjects

Our study includes three groups of subjects: (1) Healthy subjects who did not have any active medical problem and had normal blood counts; (2) Males and females with myeloproliferative neoplasms (all had well characterized myeloproliferative disorders as per World Health Organization criteria [10], namely PV, ET, and PMF) and (3) controls with reactive nonneoplastic individuals with elevated platelets wherein ET was ruled out (reactive thrombocytosis) and reactive (nonPV) polycythemias. After informed consent, 5 to 8 ml of peripheral blood was collected by venipuncture. The granulocytes fraction was isolated by gradient medium (Histopaque 1077, Sigma Aldrich) (see Section 3.1.3). This study was reviewed and approved by the institutional review board from the University of Utah (see Section 3.1.5). Table 4.1 summarizes the baseline information from subjects recruited for the study.

4.3 Methods

Confocal microspectroscopy configuration was used for hyperspectral imaging and acquisition of spectra, as described in Section 3.1.1, using 60X (NA 0.80) and 10X (NA 0.25) microscope objectives, respectively. Silver nanoparticles deposited on glass were prepared by the Tollens reaction as explained in Section 3.1.2 with a 120-second reaction time. Substrates were stored no longer than 1 day under nitrogen atmosphere from the time of preparation until their use. Granulocytes and mononuclear cell fractions were isolated by centrifugation with a density gradient medium (Section 3.1.3). Solutions of cells of at least $10^6 cells/ml$ but no more than $10^7 cells/ml$,

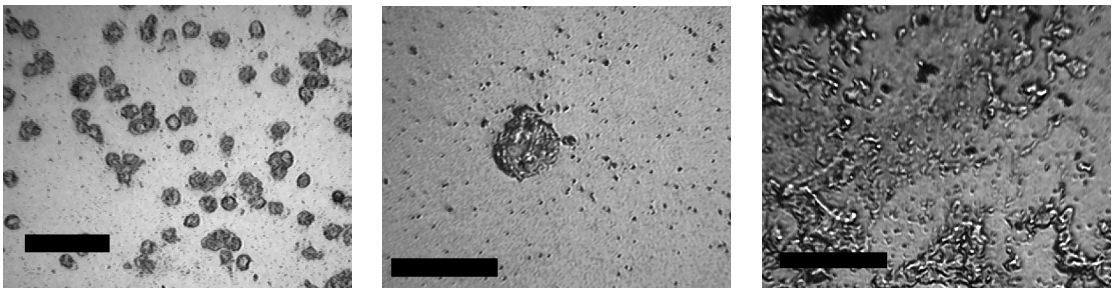
Table 4.1: Baseline characteristics of subjects enrolled in the study.

	Patients (n=75)	**Controls (n=11)**
Age (value, SD) (years)	60.8(16.0)	38.2(16.1)
Sex		
Men	28(37.3%)	6(54.5%)
Women	47(62.7%)	5(45.5%)
Disorder		
ET	25(33.3%)	
PV	29(38.7%)	
PMF	15(20%)	
Reactive, secondary	6(8%)	
JAK2 V617F mutation frequency		
ET	17/25(68%)	
PV	27/29(93.1%)	
PMF	10/15(66.7%)	
Treatment		
Hydrea	45(60%)	
Nontreated	20(26.7%)	
Anagrelide	3(4%)	
Jak2 Inhibitors	4(5.3%)	
Pegasys	2(2.7%)	
Other	2(2.7%)	
Clonality (Females)		
Clonal	26(60.5%)	
Polyclonal	5(11.6%)	
Noninformative	7(16.3%)	

were prepared in phosphate buffered saline (PBS) and stored at 4 ^{o}C until optical measurements were done, which was typically within 24 hours of preparation.

Right before the spectroscopic measurements, one or two drops of the cell solution were placed onto the silver substrate. With the help of a clean coverslip the liquid was carefully smeared and later blow-dried with nitrogen gas until most of the excess water was removed but not completely. The substrate was rinsed with distilled water to removed salt crystals left by the PBS and blowdried with nitrogen gas, again, until most of the excess water was removed but not completely. Figure 4.1 shows light microscope images of granulocytes onto silver substrates after sample preparation. Successful sample preparation would yield cells that look well defined under the light microscope (Figures 4.1A and 4.1B). Improper storage, nongentle smear and/or overdrying cause cell lysis and nonwell defined structures may be observed, as shown in Figure 4.1C. Additionally, if concentrations of cells solutions are too high, they tend to aggregate and form bigger structures where single cells are difficult to recognize.

Spectra from granulocytes and mononuclear cells from human subjects were acquired. Dark current and cosmic rays artifacts were removed from each spectrum, after which a Savitzky-Golay filter (4th Order, Zero Derivative)was applied.



A B C

Figure 4.1: Light microscope image of granulocytes onto silver nanoparticles deposited on glass with **A:** 10X magnification (scale bar= 100μm) and **B:** 60X magnification . **C:** Unsuccessful sample preparation (scale bar= 20μm)

4.3.1 Clonality studies

All studied females were evaluated for clonality of granulocytes, which was established by the genotyping of 5 X chromosome exonic polymorphisms followed by determination of the informative heterozygous X chromosome allelic mRNA usage ratio by quantitative real-time AS-PCR (qAS-PCR), as described in [11].

We previously established that exonic SNP polymorphic X chromosome allelic frequency utilization greater than 90% is an indicator of the clonal phenotype [12,13].

4.3.2 JAK2 mutational analysis

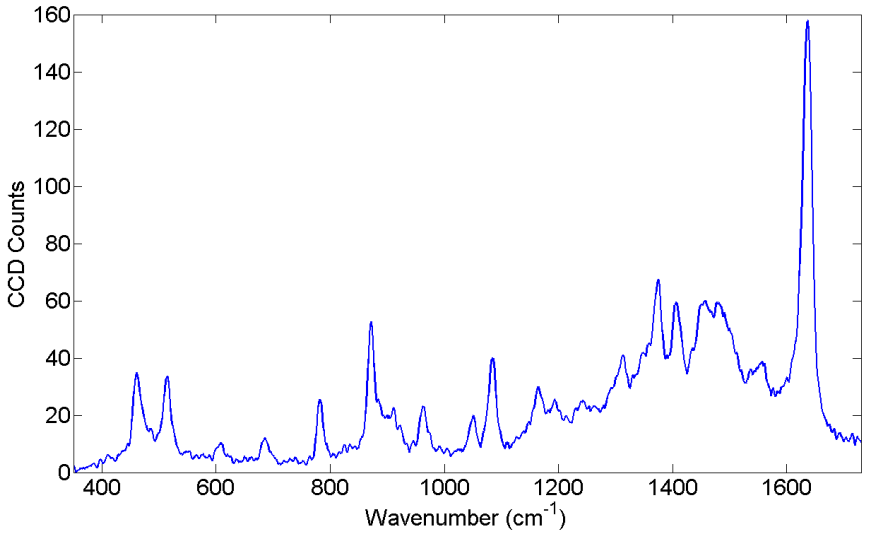
The JAK2 V617F mutational burden in separated peripheral blood cells was determined by quantitative allele specific PCR (qAS-PCR). The JAK2 V617F mutation was established by qAS-PCR utilizing mutation specific primers containing locked nucleic acid and a mismatched nucleotide [11,14].

4.4 Carbonization and silver oxides: setting up the experimental conditions

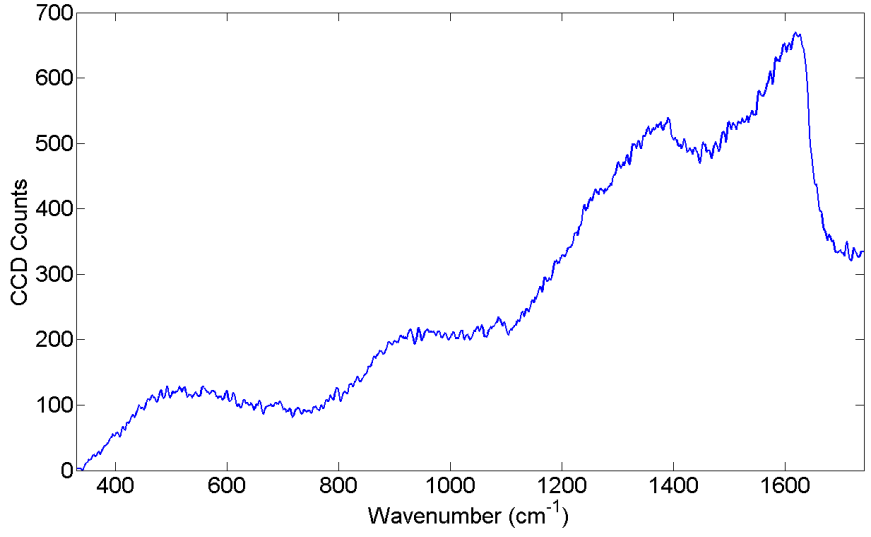
The SERS signal is often obscured by the appearance of two broad vibrational bands that are associated with the presence of amorphous carbon structures. The origin of the amorphous carbon is usually attributed to analyte decomposition caused by heat (i.e., pyrolysis or carbonization), catalytic activity of the substrate, or by contamination from carbonaceous species [15,16].

The formation of amorphous carbon from methylene blue ($C_{16}H_{18}N_3SCl$) on the Tollens substrates described in Section 3.1.2 was observed as a function of time and excitation intensity. For this study, a solution of methylene blue in ethanol was prepared and drop casted onto the substrate. The SERS spectrum of methylene blue is shown in Figure 4.2A. Figure 4.2B shows the spectrum of methylene blue after 250 seconds of exposure under 2.1 mW excitation power ($\lambda_{exc} = 488nm$). The transformation to an amorphous carbon structure is evident by the the vibrational bands at $\sim$1360 cm^{-1} and $\sim$1580 cm^{-1}. The bands at $\sim$500 cm^{-1} and $\sim$900 cm^{-1} are assigned to oxidation of the substrate, specifically AgO and Ag_2O [17,18].

The dynamics of the carbonization depend upon laser excitation power. Examples



A



B

Figure 4.2: SERS spectrum of methylene blue ($C_{16}H_{18}N_3SCl$) **A:** on Tollens substrate with laser wavelength 488nm (0.24 mW, 60X microscope objective) and **B:** after 250 seconds of exposure under 2.1 mW excitation power.

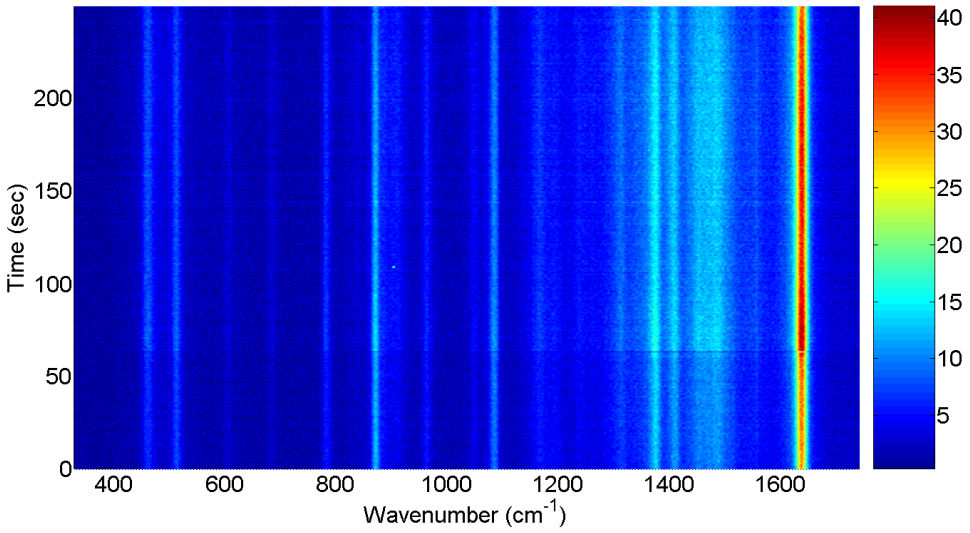
of this are shown in Figures 4.3A and 4.3B with excitation power of 0.24 mW and 2.1 mW, respectively. In Figure 4.3A, the spectrum does not change in time, whereas in Figure 4.3B, the appearance of the spectral features described above is observed. Figure 4.4A shows the time dependence of the total emission intensity (TEI), normalized to t=0, for various excitation intensities, which has been fitted by an exponential of the form $TEI(t) = T_0 exp(-t/\tau) + c$. The time parameter τ is plotted in Figure 4.4B as function of excitation intensity. Only with very low intensity ($\leq 0.21mW$) the formation of amorphous carbon is not observed.

SERS spectra from granulocytes exhibit a similar time dependence as observed in methylene blue with a very important difference. There is an initial period of time of about 10 seconds for which the spectra is unchanged, as shown in Figure 4.5. Measurements on cells are limited within that period of time. Thus, the integration time per accumulation was set to 0.1 seconds and 10 accumulations per spectra were acquired, making the measurement about 1 second long.

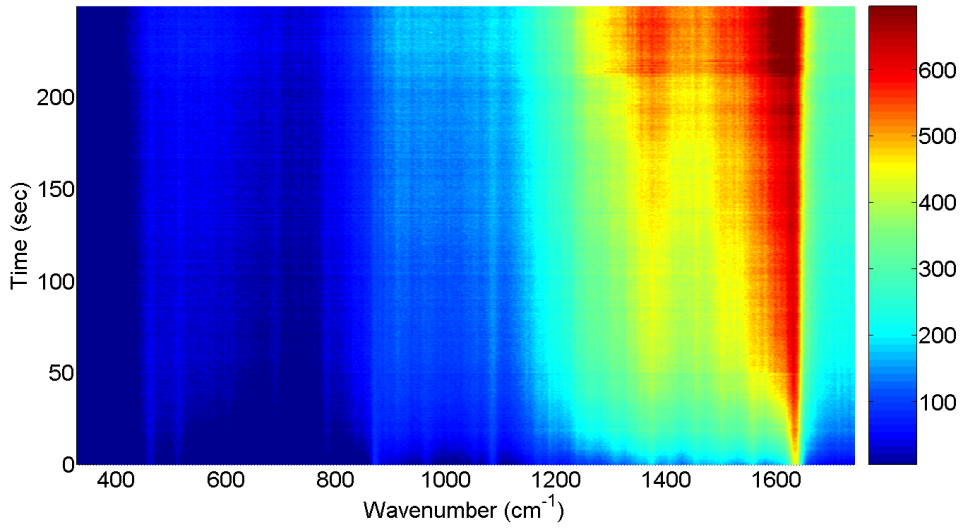
Contamination by carbonaceous species is one of the major problems encountered when doing SERS [15, 16, 19]. Some authors even claim the presence of amorphous carbon to be ubiquitous in SERS, especially if silver substrates are used [16, 19]. Protocols aimed to clean the substrates have been largely studied with various rates of success [20]. It was determined by energy dispersive X-ray spectroscopy (EDS, 10kV accelerating voltage)[3] that UV-ozone cleaning protocol did not reduced the amount of carbon present on the substrate, but it increased the amount of oxygen (see Table 4.2). The increased oxygen amount was attributed to oxidation of the silver induced by the UV radiation [21]. Given these results, a cleaning protocol was not adopted; instead the substrates were freshly prepared before measurements and used within a couple of hours of the preparation.

The presence of the two broad vibrational bands associated with silver oxide described above was also observed when measurements were performed under nitrogen atmosphere, suggesting that silver oxide was present before the measurements and

[3]Scanning electron microscope, FEI NanoNova FEG-SEM 630, Hillsboro, OR USA.

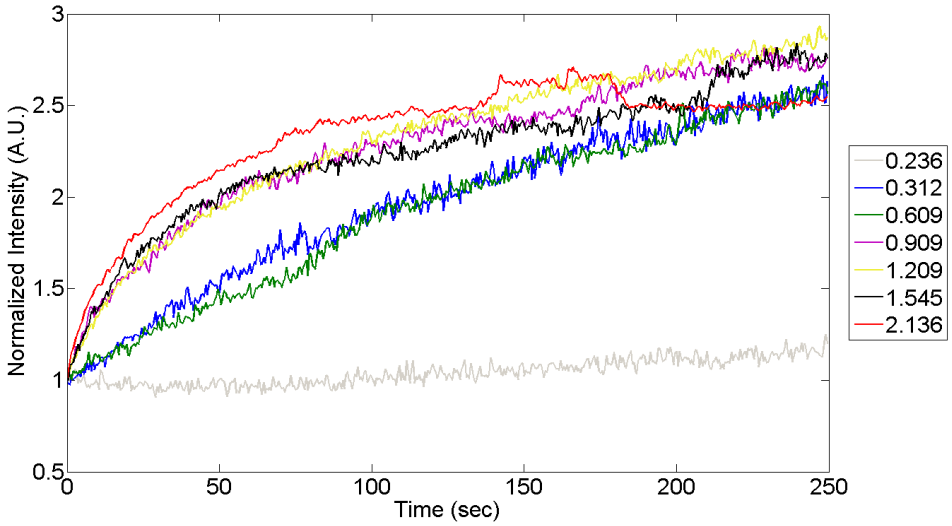


A

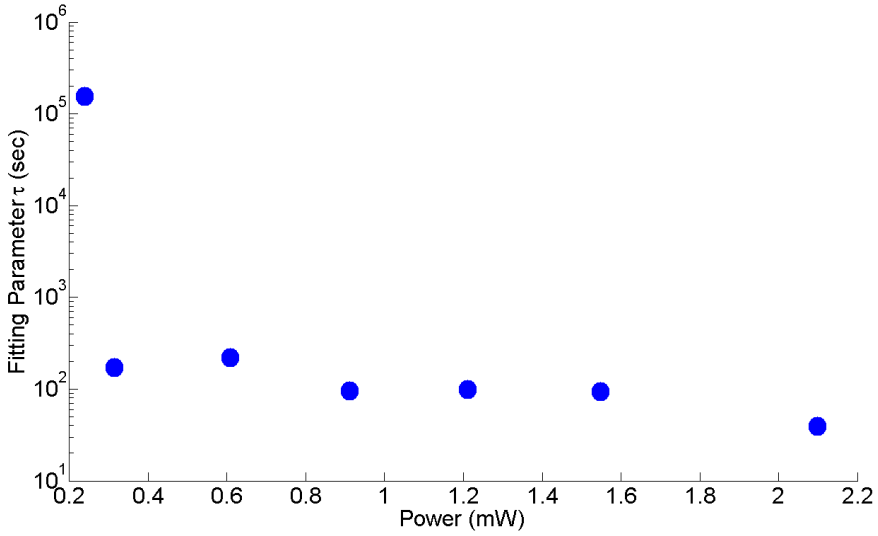


B

Figure 4.3: SERS spectrum of methylene blue as a function of time under **A:** 0.24mW excitation power and **B:** 2.1mW excitation power.



A



B

Figure 4.4: Dynamics of the SERS spectrum of methylene blue. **A:** Total emission intensity, normalized to t=0, as a function of exposure time for various excitation intensities (shown in mW). **B:** Each curve in A was fitted to an exponential of the form $TEI(t) = T_0 exp(-t/\tau) + c$. Fitting parameter τ is shown as a function of excitation intensity.

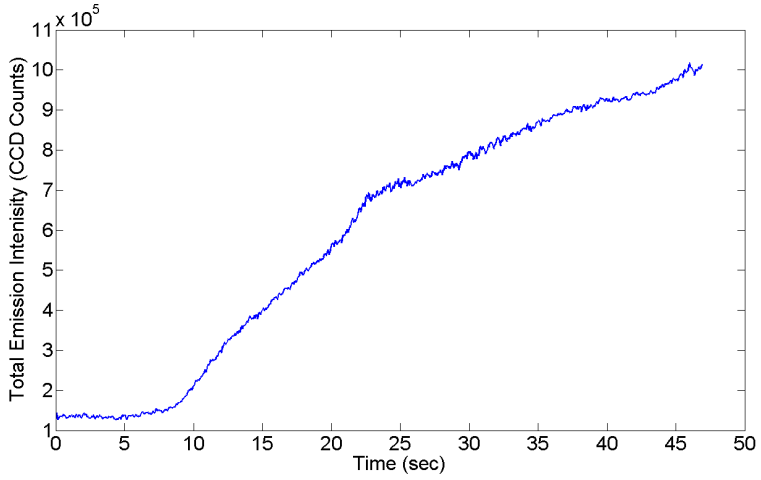


Figure 4.5: Typical total emission intensity as a function of time from a human granulocyte lying on top of a Tollens substrate(10X Microscope Objective, 3.5mW excitation power and 488nm excitation wavelength)

Table 4.2: Weight percentage as measured by EDS of Tollens substrates before and after 30 seconds of UV-Ozone exposure.

Element	No UV-Ozone ($Wt\%$)	After UV-Ozone ($Wt\%$)
Silver	37.86	30.3
Aluminium	1.02	1.43
Boron	25.79	25.53
Carbon	5.52	5.01
Potassium	2.34	2.62
Sodium	2.46	2.99
Oxygen	12.19	17.06
Silicon	12.81	14.99

probably formed immediately after the Tollens reaction took place. This fact made the use of a nitrogen atmosphere ineffectual as it did not prevent the formation of silver oxides. Therefore, the SERS measurements were performed under normal atmospheric conditions.

4.5 SERS on human granulocytes: results

The emission spectrum was measured at the single cell level under the conditions described above. Figure 4.6 shows the average spectra for each of the four groups, namely ET, PV, PMF and Control. The bands centered at $\sim$500 cm^{-1} and $\sim$1000 cm^{-1} are assigned to AgO and Ag_2O in the substrate [17, 18]. The emission spectra also show the formation of an amorphous carbon structure, as evident by the vibrational bands at $\sim$1370 cm^{-1}, also known as the D-band, and $\sim$1600 cm^{-1}, also known as the G-band [18, 22, 23]. In addition, the spectra also contain a broad background emission in the range of 500-4500 cm^{-1} discussed in more detail below.

Quantification of spectral differences between affected and healthy subjects was done by statistical analysis of the total emission intensity (after normalization to the vibrational band centered at $\sim$500 cm^{-1}), as well as ratios of integrated intensity over paired scattering frequency intervals. For each subject, average values from all granulocytes were calculated. For each defined quantity, ratio and total integrated emission, *normality* and *homoscedasticity* were evaluated using the modified Kolmogorov-Smirnov and the Levene test, respectively. Significant differences between controls and patients were calculated by the nonparametric Kruskal-Wallis test or the student's t-test for unequal variances, as necessary. For ratios of intensity, significant level was adjusted to multiple comparisons (see Section 3.2.1). P-values below 0.0004 were considered significant; for any other variable, the significant level was set to 0.05.

We found that the TEI is statistically significantly higher in control subjects than in MPN patients (p-value$<$ 0.0014) and in reactive nonneoplastic subjects (p-value$<$ 0.05), but not among all three tested MPNs (Figure 4.7). We identified three viable intensity ratios that displayed significant differences not only between controls

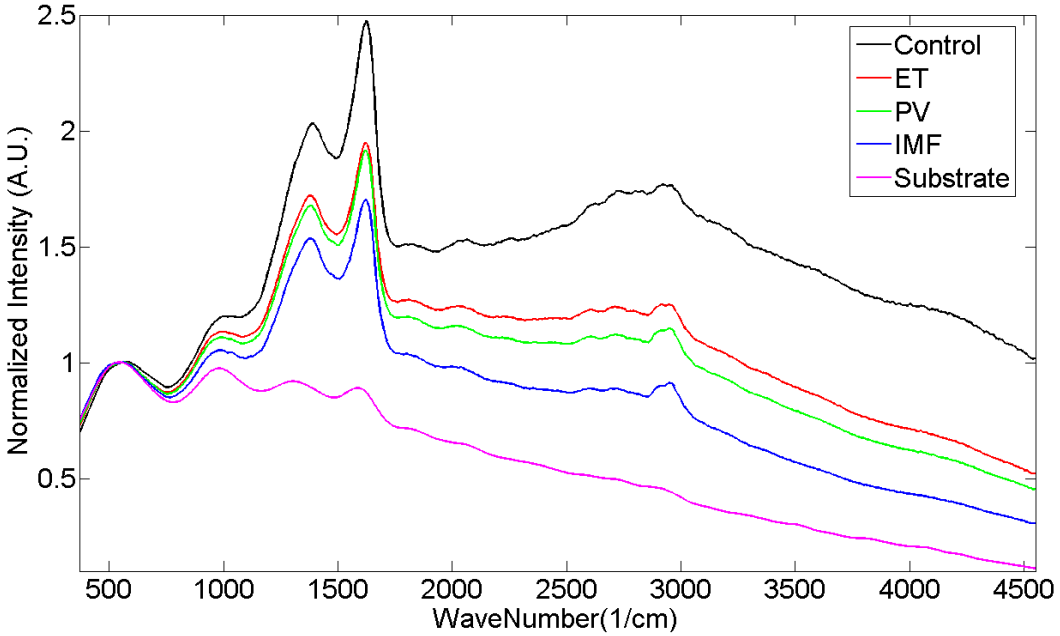


Figure 4.6: Average emission spectra from granulocytes from all controls (black), ET (red), PV (green) and PMF (blue) patients. The emission spectrum from the substrate is shown in magenta. The emission spectrum from the substrate is shown in magenta. The emission spectra were normalized to the Raman peak at $\sim$500 cm^{-1} that originates from silver oxides in the substrate. The broad road emission bands around 1600 and 1370 cm^{-1} correspond to the D and the G band, respectively, from amorphous carbon [18, 22, 23].

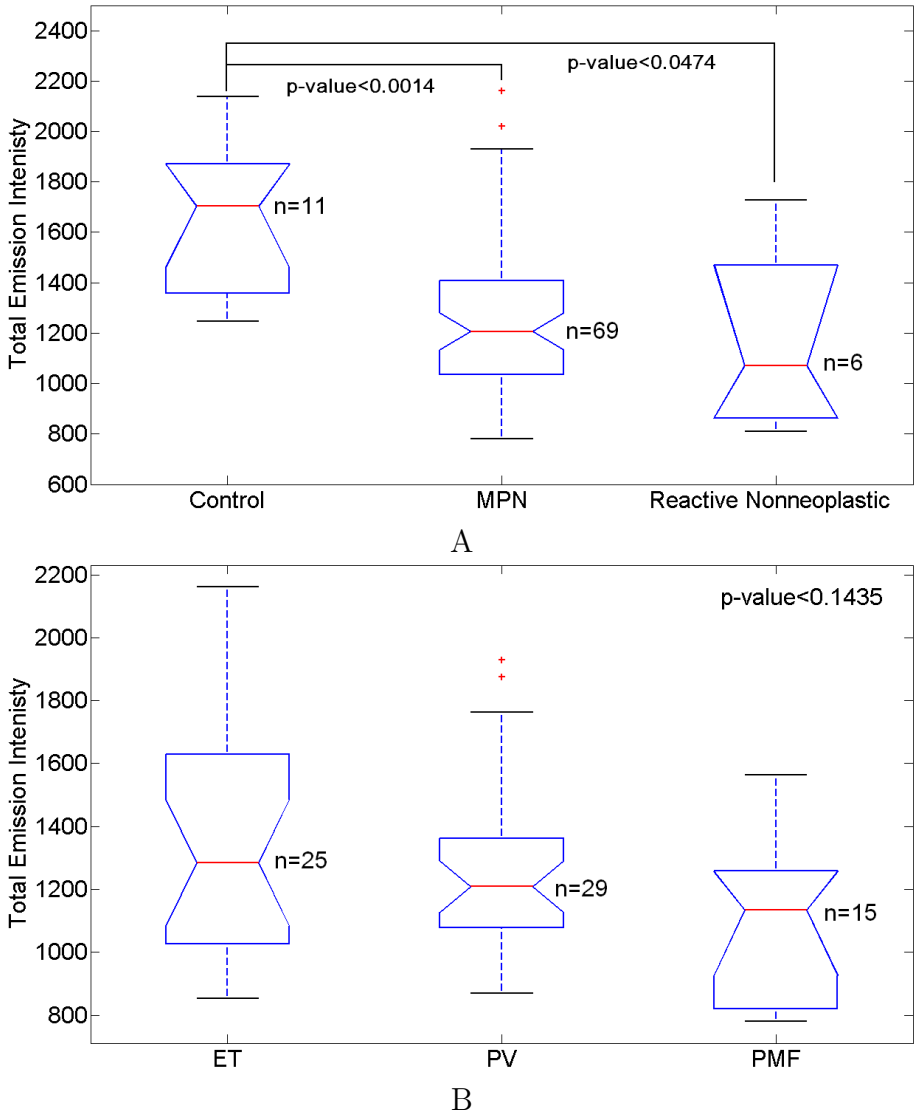


Figure 4.7: Box plots of **A:** normalized total emission intensity for controls, MPN patients and reactive nonneoplastic subjects, respectively; and **B:** normalized integrated emission intensity for each MPN; the accompanying numbers denote the number of specimens. The top and bottom boxes correspond to the first and third quartile, respectively. The whiskers lengths are 1.5 times the interquartile range. P-value is the probability that the two populations come from distributions with the same mean ratio. Crosses are outliers.

and MPN patients, but also between ET and the other MPN patients (see Table 4.3). This is of particular importance because ET is largely diagnosed by exclusion. A total of 17 different emission intensity ratios were identified that showed significant discrimination between patients and healthy subjects (Table 4.4). These 17 intensity ratios come from six frequency ranges paired in three groups (see Figure 4.8).

We explored the potential of the total integrated intensity and intensity ratios obtained from SERS spectra of granulocytes as markers for MPNs. As described in Section 3.2.2, tests with areas under the *receiver operating characteristic* (ROC) curve larger than 0.8 are considered very good, whereas tests with areas under the curve (AUC) larger than 0.9 are considered excellent. The ROC curve indicates that the total integrated emission intensity is a good test for the presence of an MPN with an AUC of 0.818 (95% CI: 0.707 - 0.930). See Figure 4.9. At the cut-off point, sensitivity is 82.3% (95% CI: 72.7% - 91.8%) and specificity is 72.7% (95% CI: 46.4% - 99.0%). ROC analysis for all intensity ratios that exhibit a significant difference between controls and MPN patients (p-value$<$ 0.0004) is shown in Table 4.4, where AUC, sensitivity and specificity at the cut-off point are listed. The ROC curves reveal that intensity ratios in Table 4.4 are excellent tests for the presence of an MPN, as they have an AUC larger than 0.9, and provide higher specificity and sensitivity than the total integrated intensity. The ROC curves of the best performing ratios from each of the frequency intervals pairs are shown in Figures 4.10 and 4.11. Although some intensity ratios show a significant difference among ET and the other

Table 4.3: P-values as found by Kruskal-Wallis test. Groups are defined as MPN patients versus controls (P vs C) and ET patients versus both PV and PMF patients (ET vs PV,PMF).

Pair	**Frequency Interval 1** (cm^{-1})	**Frequency Interval 2** (cm^{-1})	**p-value P vs C**	**p-value ET vs PV, PMF**
15	1384 - 1863	3238 - 3624	0.0002	0.0409
16	1138 - 1674	2833 - 3078	0.0004	0.0367
17	1148 - 1674	2229 - 2579	0.0162	0.0354

Table 4.4: Receiver operating characteristic (ROC) analysis for ratios from paired frequency intervals. AUC= Area under the ROC curve, SE=Standard Error. Sensitivity and specificity are calculated at the cut-off point, defined as the closest point to the perfect test on the ROC curve. The best ratio is highlighted in red.

Pair	**Frequency Interval 1** (cm^{-1})	**Frequency Interval 2** (cm^{-1})	**AUC (SE)**	**AUC (95% CI)**	**Sensitivity (95% CI)**	**Specificity (95% CI)**
1	1769-1821	4134-4515	0.912 (0.035)	0.842-0.982	79.0 (68.9-89.2)	100 (100-100)
2	1863-1914	4204-4448	0.911 (0.036)	0.840-0.981	88.7 (80.8-96.6)	81.8 (59.0-100)
3	1954-2007	4204-4380	0.911 (0.036)	0.840-0.981	88.7 (80.8-96.6)	81.8 (59.0-100)
4	2048-2098	4273-4380	0.909 (0.036)	0.838-0.980	88.7 (80.8-96.6)	81.8 (59.0-100)
5	2139-2188	4273-4380	0.909 (0.036)	0.838-0.980	88.7 (80.8-96.6)	81.8 (59.0-100)
6	2139-2188	4478-4558	0.909 (0.036)	0.838-0.980	88.7 (80.8-96.6)	81.8 (59.0-100)
7	1769-1914	4410-4448	0.909 (0.036)	0.838-0.980	88.7 (80.8-96.6)	81.8 (59.0-100)
8	1769-2188	4342-4380	0.911 (0.036)	0.840-0.981	88.7 (80.8-96.6)	81.8 (59.0-100)
9	1769-2188	4273-4311	0.912 (0.035)	0.843-0.982	88.7 (80.8-96.6)	81.8 (59.0-100)
10	1769-2098	4204-4242	0.911 (0.036)	0.840-0.981	88.7 (80.8-96.6)	81.8 (59.0-100)
11	3940-3981	4134-4242	0.913 (0.035)	0.845-0.982	87.1 (78.8-95.4)	81.8 (59.0-100)
12	4063-4102	4134-4380	0.922 (0.033)	0.858-0.986	83.9 (74.7-93.0)	90.9 (73.9-100)
13	3920-4102	4342-4380	0.912 (0.035)	0.842-0.982	80.7 (70.8-90.5)	90.9 (73.9-100)
14	3773-4173	4204-4242	0.915 (0.035)	0.847-0.983	83.9 (74.7-93.0)	90.9 (73.9-100)
15	1384-1863	3238-3624	0.909 (0.036)	0.838-0.980	88.7 (80.8-96.6)	81.8 (59.0-100)

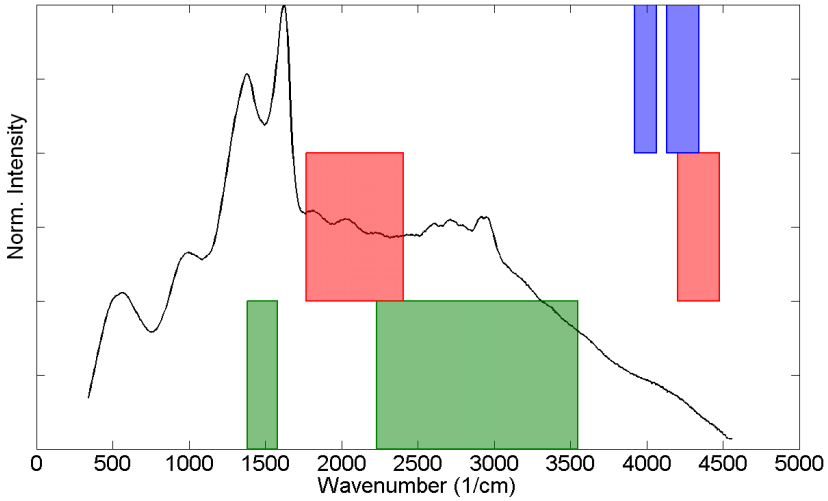


Figure 4.8: Paired areas from where ratios of integrated emission intensities showed statistically significant difference between control group and patients (blue, red and green) and between ET patients and PV,PMF patients (green).

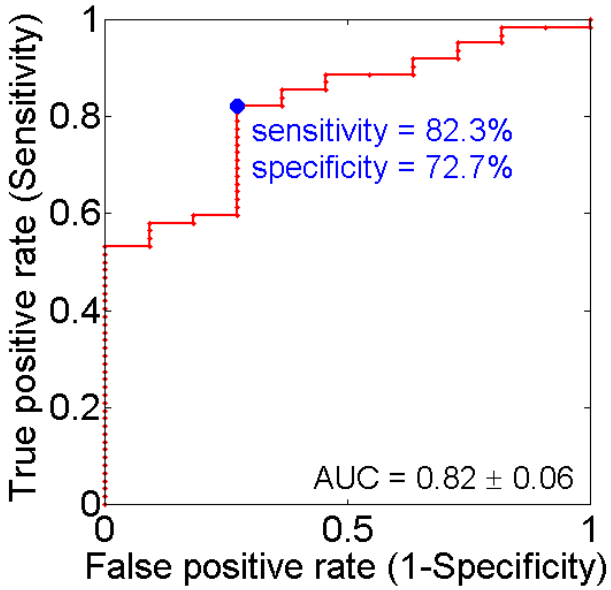


Figure 4.9: ROC (Receiver operating characteristic) curve for the total emission intensity. Blue dot is the cut-off point defined as the closest to the ideal test point. At cut off point sensitivity is 83.1% (95% CI: 73.5% - 92.6%) and specificity is 72.7% (95% CI: 46.4% - 99.0%). Negative=Control, Positive=MPN patient. AUC=Area under the curve.

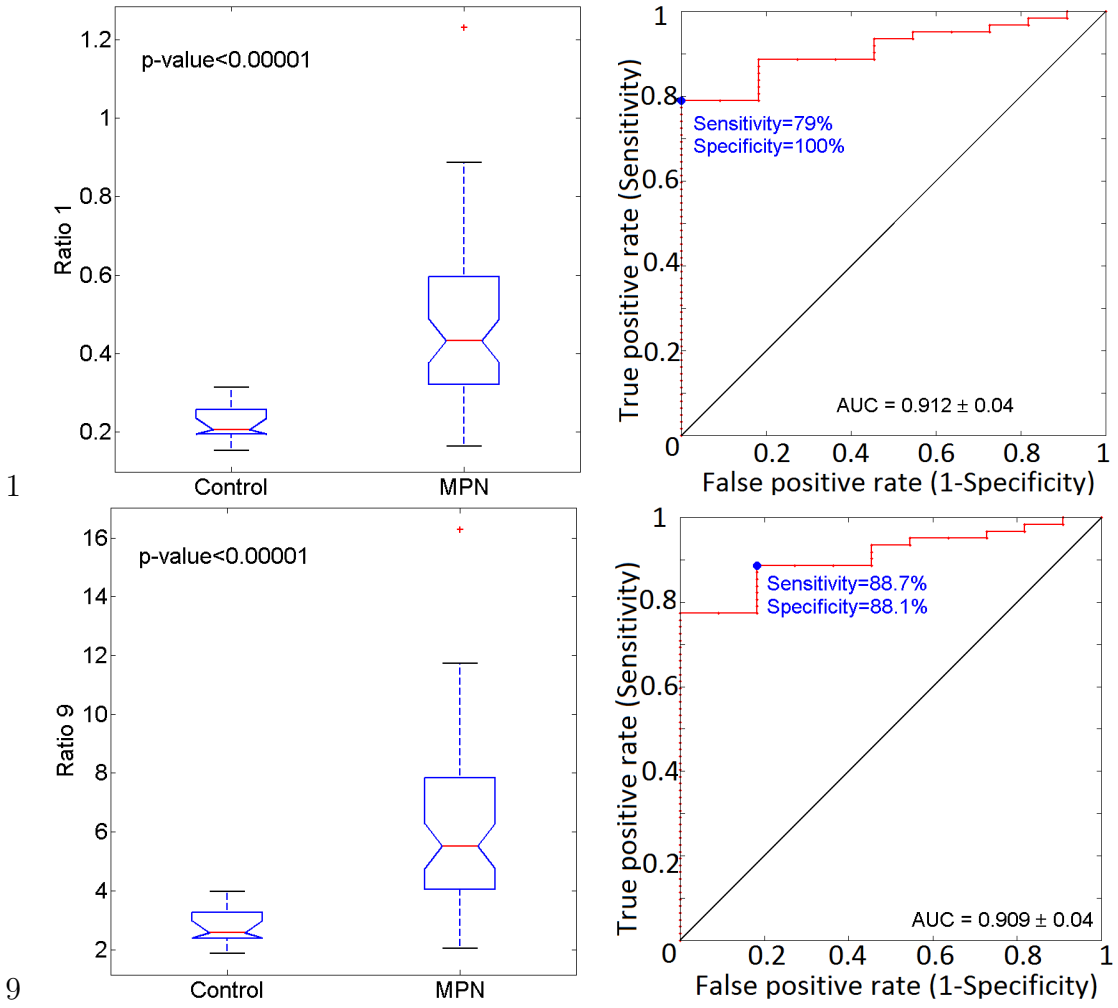
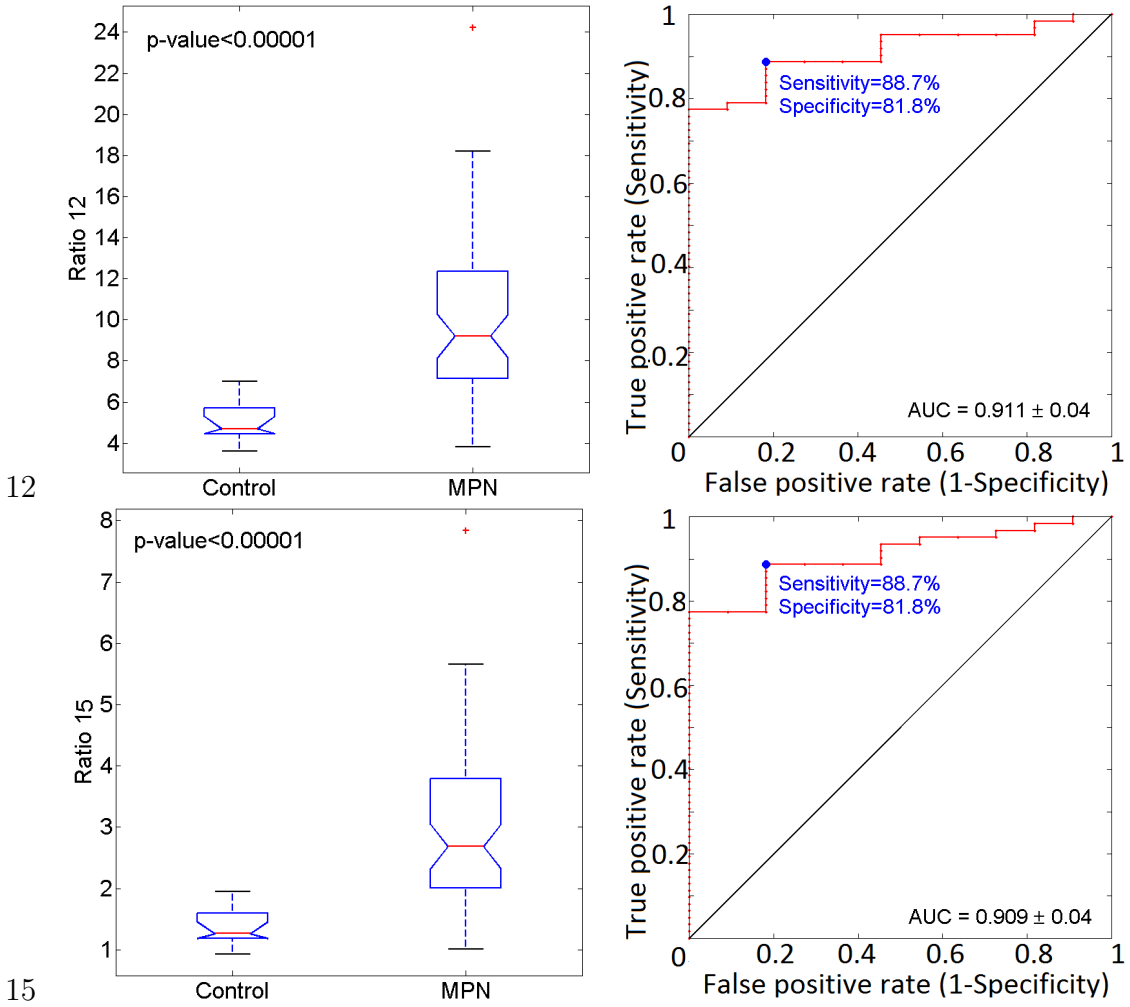


Figure 4.10: Statistical analysis for best performing ratios **Left Column:** Ratio identification number as defined in Table 4.4. **Middle column:** Ratios of integrated intensity for controls and MPN patients. The top and bottom boxes correspond to the first and third quartile, respectively. The whiskers lengths are 1.5 times the interquartile range. P-value is the probability that the two populations come from distributions with the same mean ratio. Crosses are outliers. **Right Column:** ROC curve for ratios of intensity. Blue dot is the cut-off point defined as the closest to the ideal test point. Sensitivity and specificity are shown at cut-off point. Negative=Control, Positive=MPN patient. AUC=Area under the curve.



12

15

Figure 4.11: Statistical analysis for best performing ratios **Left Column:** Ratio identification number as defined in Table 4.4. **Middle column:** Ratios of integrated intensity for controls and MPN patients. The top and bottom boxes correspond to the second and third quartile, respectively. The whiskers lengths are 1.5 times the interquartile range. P-value is the probability that the two populations come from distributions with the same mean ratio. Crosses are outliers. **Right Column:** ROC (Receiver operating characteristic) curve for ratios of intensity. Blue dot is the cut-off point defined as the closest to the ideal test point. Sensitivity and specificity are shown at cut-off point. Negative=Control, Positive=MPN patient. AUC=Area under the curve.

two disorders (p-value< 0.05), when the ROC curve was evaluated with ET patients as positive cases and PV and PMF patients as negative, a diagnostic value was not found (AUC< 0.7).

Figure 4.12 shows the cross correlation coefficients between ratios of intensity, total emission intensity and clinically relevant variables (age and *JAK2*V617F allelic burden) for all subjects. Correlations between *JAK2*V617F allelic burden and total emission intensity or spectral ratios were not found. A slight correlation between age and spectroscopic variables was found (smaller that 0.41). Correlations between ratios and TEI are moderate (less than 0.8), while correlations between ratios 1-14 are very strong (larger than 0.8). Ratio 17 (see Table 4.3) is weakly correlated to all of the other variables (< 0.65) except for ratios 15 and 16 with which it is highly correlated (> 0.89). High correlations among ratios 1 to 14 agree with them having the same origin in the broad continuum.

The surfaced enhanced emission spectra from granulocytes are composed of two parts: (*i*) a broad luminescent background (also known as the continuum spectrum in the literature), and (*ii*) Raman scattering vibrational bands associated with vibrations of amorphous carbon and silver (I, II) oxide. In order to separate the two emission components, we used the algorithm described by Zhao *et al* [24], with a third order polynomial. Figure 4.13A shows the average spectra of each of the populations studied. Figure 4.13B shows the background continuum component, and Figure 4.13C shows the vibrational Raman scattering contribution of each of the populations. The continuum spectrum of the patients displays a blue shift compared to that of controls, as indicated by the spectral position of the maximum (Figure 4.13D); the continuum spectrum for ET patients shows the least shift, whereas the spectrum from PMF patients shows the most shift.

A control experiment was performed for studying the underlying mechanism for the spectral shift obtained in MPN patients. We incubated granulocytes with various concentrations of 2-deoxy-D-glucose (2DG), which is a glucose molecule that is phosphorylated by glucose hexokinase into 2-deoxyglucose-6-phosphate (2D6G) and thus cannot undergo further glycolysis. 2D6G accumulates in the cell and may serve

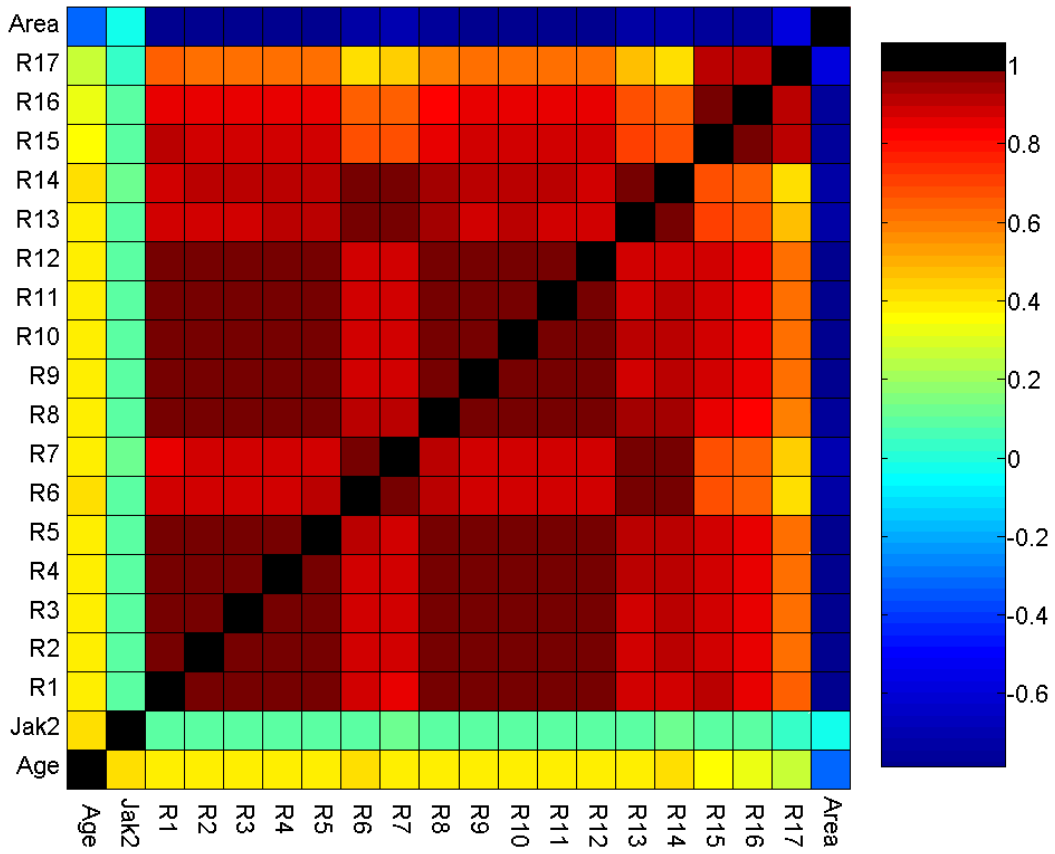


Figure 4.12: Crosscorrelation matrix for ratios of intensity, total emission intensity and clinically relevant variables, age and *JAK2*V617F allelic burden for all subjects recruited in the study.

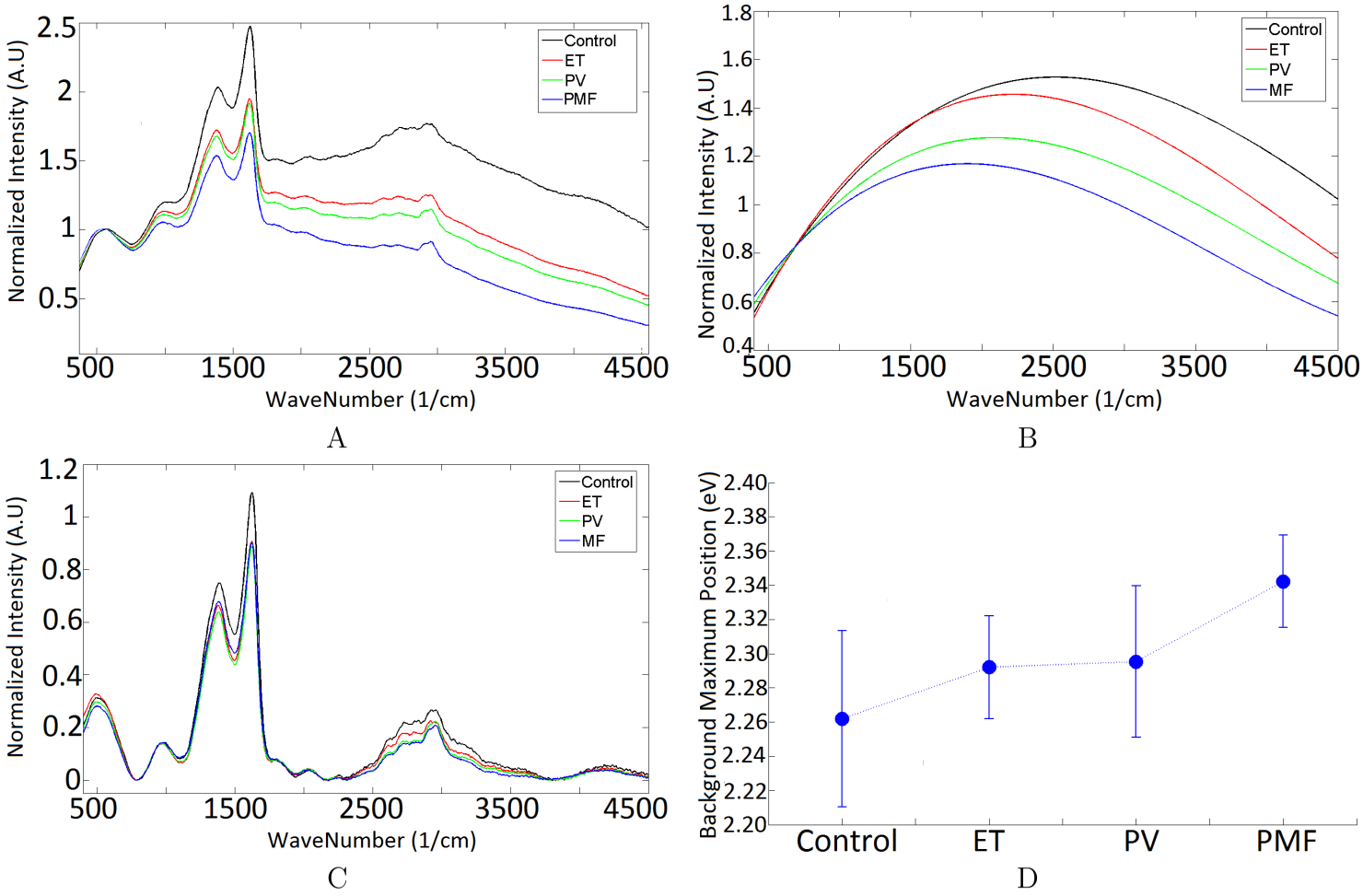


Figure 4.13: SERS spectra from granulocytes. **A:** Average spectra from all controls (black), ET (red), PV (green) and PMF (blue) patients. The spectrum from the substrate is shown in magenta. Spectra were normalized to peak at ~500 cm^{-1} from silver oxides in the substrate. Broad emission bands around 1370 and 1600 cm^{-1} corresponding go the D and the G bands, respectively [18, 22, 23]. **B:** Background of the spectra in A fitted to a third order polynomial. **C:** Vibration related Raman scattering contribution extracted from the spectra in A. **D:** The peak photon energy of the background emission for controls, ET, PV and PMF patients.

a measure of cellular glucose uptake [25].

Granulocytes isolated from a healthy subject were incubated for 15 minutes at 37^{o}C with 2DG solutions in PBS. Briefly, after isolation, cells solution was divided into four identical aliquots and centrifuged at 400 g for 10 min. Pellets were resuspended in 1 ml of 1 mM, 5 mM and 10 mM 2DG solutions. The fourth aliquot was resuspended and incubated in 1ml of PBS without 2DG to be used as a control. Optical measurements were performed within 4 hours of incubation. SERS spectra were obtained and postprocessed under the conditions described above (Section 4.3).

We found that the emission background contribution to the SERS spectra from granulocyte is blue shifted with increasing concentrations of 2DG (Figure 4.14D). This suggests that the blue shift in the background continuum spectrum may be explained in part by changes induced by higher glucose uptake in the MPN patients' cells.

4.5.1 Time resolved spectroscopy from granulocytes

We believe the broad continuum emission observed in the granulocytes spectra from healthy subjects and patients is due to a photoluminescent process. To corroborate this statement, time resolved measurements of the emission from granulocytes on the Tollens substrate were acquired using the second harmonic from a tunable Ti:Sapphire laser (140 fs pulse duration, 80 MHz repetition rate) and a streak camera ($\sim$5 ps time resolution) synchronized to the excitation pulse train. The sample was illuminated at a 30 ^{o} angle (with respect to the normal), and collection was done in the normal direction with a microscope objective (NA 0.55). After appropriate filtering of the excitation light, the emission was dispersed using a 50 grooves/mm grating (600 nm blaze wavelength) onto the streak camera. The excitation power was set to $\sim$10 mW and two excitation wavelengths were used, 488 nm and 458 nm, respectively. Additionally, long exposure/long integration time measurements (Int Time $\sim$2 sec) were done using 488 nm and 458 nm excitation wavelength from the Ti:Sapphire laser; collected light was dispersed onto a CCD camera (thermoelectrically cooled to -30 ^{o}C) using a 150 grooves/mm grating (500 nm blaze wavelength).

Long exposure measurements confirmed that the broad bands observed in the

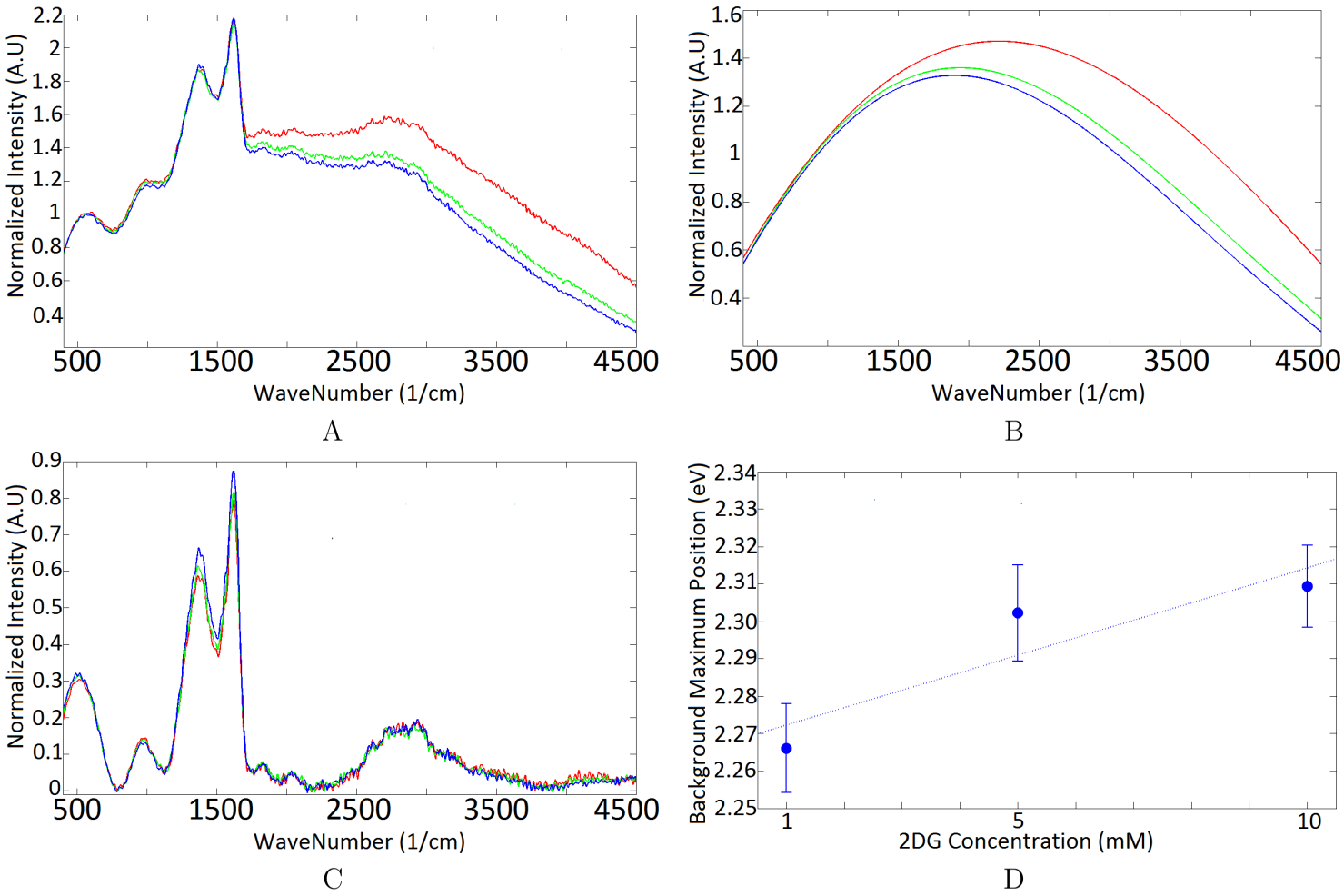


Figure 4.14: SERS spectra from granulocytes incubated with 2DG. **A:** Spectra from granulocytes incubated with 1 mM (red), 5 mM (green) and 10 mM (blue) of 2DG for 15 minutes at 37 °C. **B:** Background of spectra in A fitted to a third order polynomial. **C:** Spectra in A minus the background. **D:** Photon energy of the background peak position of the spectra in B.

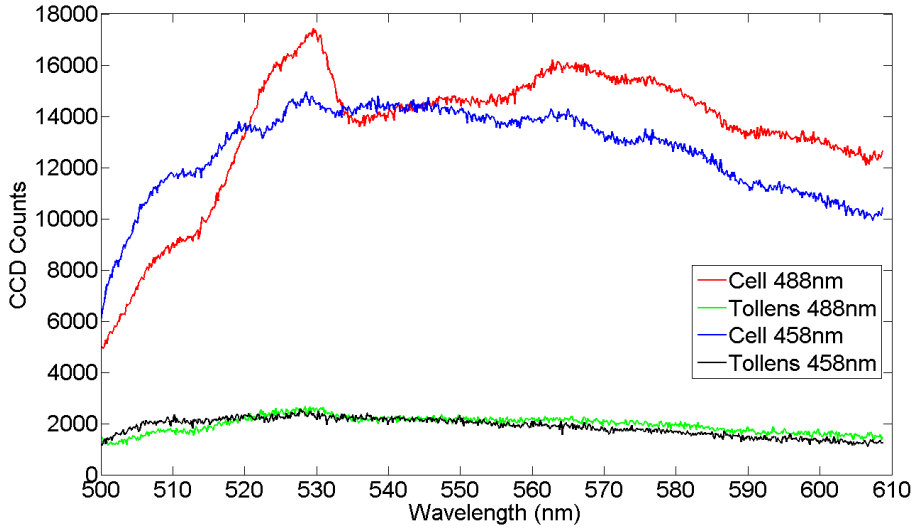
1300-1600 cm^{-1}, when using 488 nm excitation wavelength, are due to vibrational raman scattering. As observed in Figure 4.15, the relatively sharper features in the 510-540 nm region from the cell's spectrum are observed under 488 nm but not under 458nm excitation wavelength. On the other hand, the broad continuum spectrum is still present, confirming its nonvibrational origin. Additionally, the emission spectrum from the substrate *(i)* has lower total emission when compare to that for the cell (lying on top of the substrate), and *(ii)* displays no vibrational bands, suggesting that most of both components, the luminescent and the vibrational contributions to the measured spectra, arise from the cell or the cell-metal interface.

Further confirmation of the luminescent origin of the broad continuum spectrum is given by the time resolved measurements done on a granulocyte shown in Figure 4.16A. The peaks shifted from the excitation wavelength appeared only in the instantaneous spectrum, expected for raman bands, whereas the delayed emission spectrum displays the broad continuum (Figure 4.16B) [26]. Moreover, the interval 509 to 528 nm (equivalent to $\sim$1000 to 1500 cm^{-1}) showed a faster decay than the rest of the spectrum (Figure 4.17).

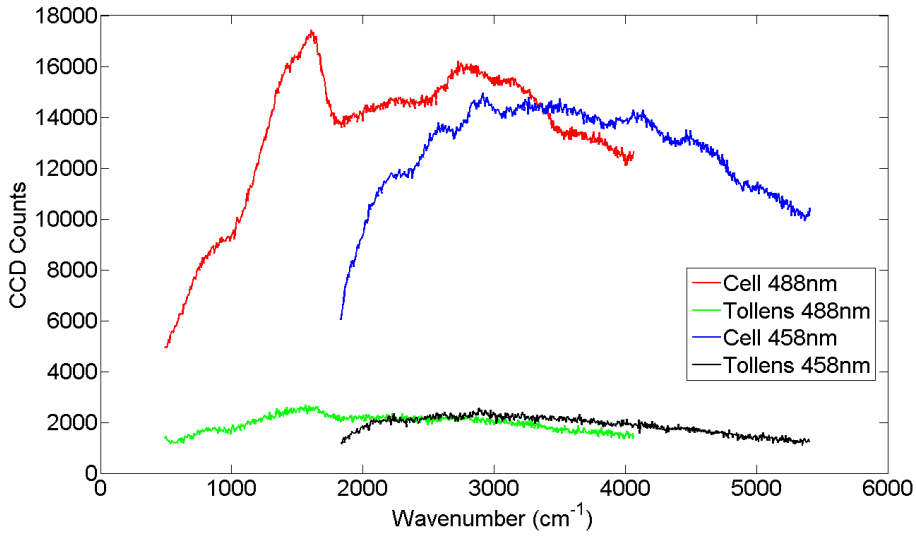
4.6 Expression of Hypoxia Induced Factor targets

Hypoxia Induced Factor (HIF) expression is induced in a variety of cancer cells due to neoplastic metabolic alteration or as an adaptive response that leads to augmentation of aerobic glycolysis or Warburg effect in cancer tissues. This transformation results in cancer cells consuming more glucose under aerobic conditions caused by upregulation of HIF transcription factor [9].

Under the hypothesis that MPNs may display Warburg effect we investigated the expression of hypoxia induced factor-1 (HIF-1) targets: SLC2A1 (glucose transporter 1) and PDK1 (pyruvate dehydrogenase kinase, isozyme 1) [27] and hypoxia induced factor-2 (HIF-2) and HIF-1 target VEGF-A (vascular endothelial growth factor A) [28]. SLC2A1 gene codes for a major glucose transporter GLUT1 normally expressed in erythrocytes and endothelial cells [29]. It has in been shown to be upregulated in malignant epithelial tissue and its expression has been associated with aggressiveness,

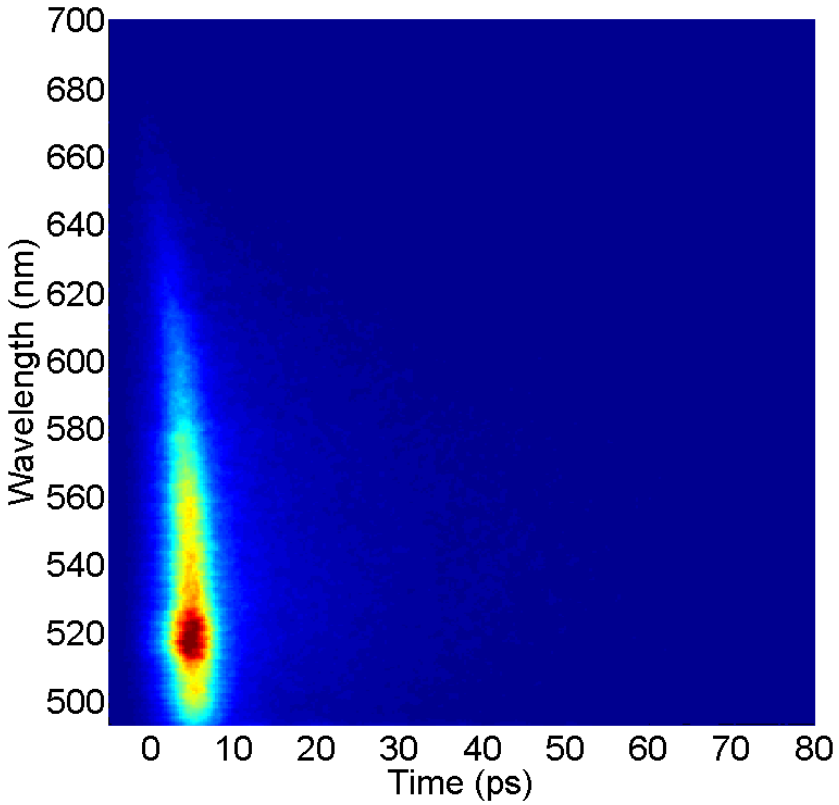


A

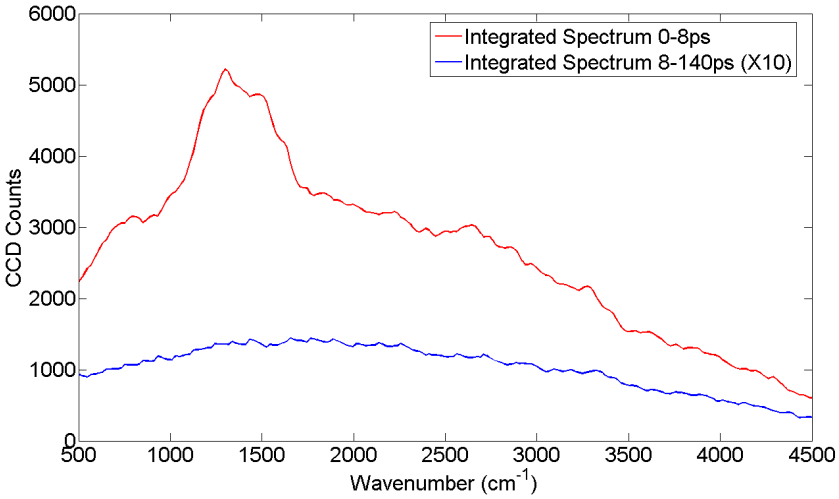


B

Figure 4.15: Long exposure/long integration time spectra of a granulocyte onto silver nanoparticles substrate and substrate itself, acquired using 488 nm and 458 nm ps excitation wavelengths **A:** as a function of wavelength and **B:** as a function of photon energy from the excitation expressed in wavenumber.



A



B

Figure 4.16: Time dependence measurements on granulocytes. **A:** Time resolved spectra from a granulocyte onto Tollens substrate generated with 488nm pulsed excitation wavelength. **B:** Integrated instantaneous (0 to 8ps) and delayed (8 to 140 ps) spectra.

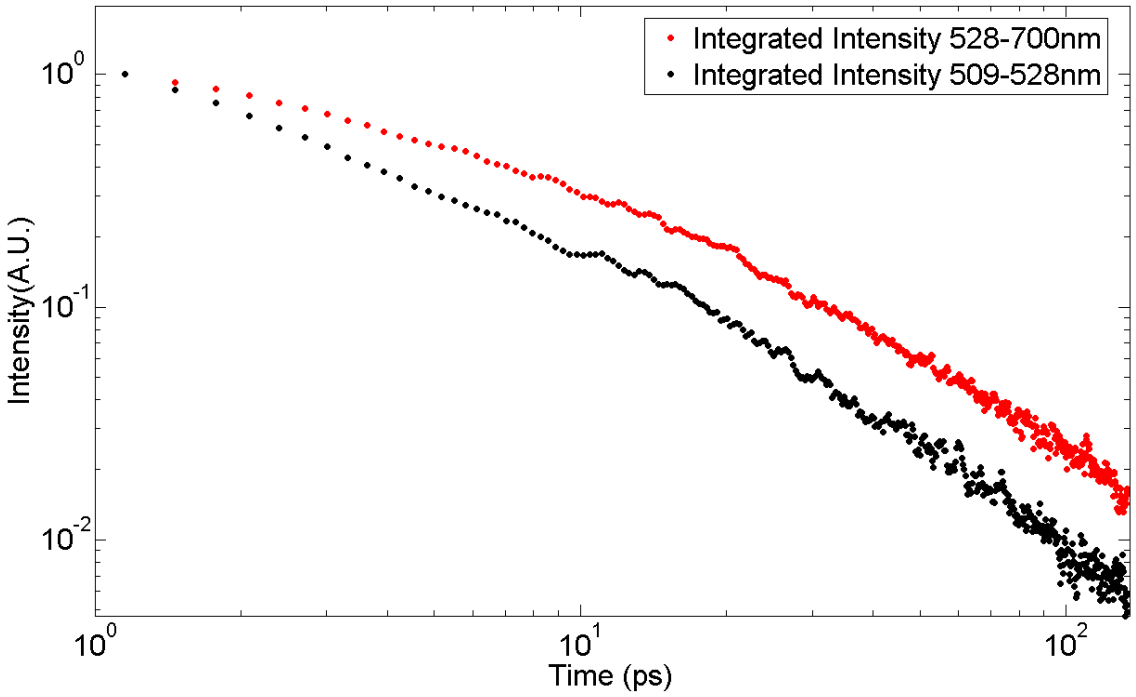


Figure 4.17: Time dependence measurements on granulocytes. Normalized intensity decay of two wavelength intervals in the spectrum, plotted in logarithmic scale.

invasiveness and poor prognosis in various neoplasms [30–33]. PDK1 genes codes an isozome of the pyruvate dehydrogenase kinase (PDK). The role of PDK is an enzyme that phosphorylates pyruvate dehydrogenase, which in turn inactivates it. Pyruvate dehydrogenase is one of the enzymes that make up part of the pyruvate dehydrogenase complex, which plays an important role in the conversion of pyruvate into Acetyl-CoA (see Section 2.3.2). PDK1 has been shown to be critical for the attenuation of the production of reactive oxygen species in the mitochondria, maintenance of ATP level and adaptation to hypoxia [27]. Finally, VEGF-A codes for a mitogen protein that promotes the growth of new blood vessels.

We did not find a significant difference in VEGF-A expression in MPN patients when compared with the control group. In contrast, we found that SLC2A1 is upregulated in MPN patients, whereas PDK1 is downregulated (Figure 4.18). Dysregulation of SLC2A1 and PDK1, both of which are along glycolytic pathways, could in turn account for higher glucose uptake and changes in the rates at which pyruvate is

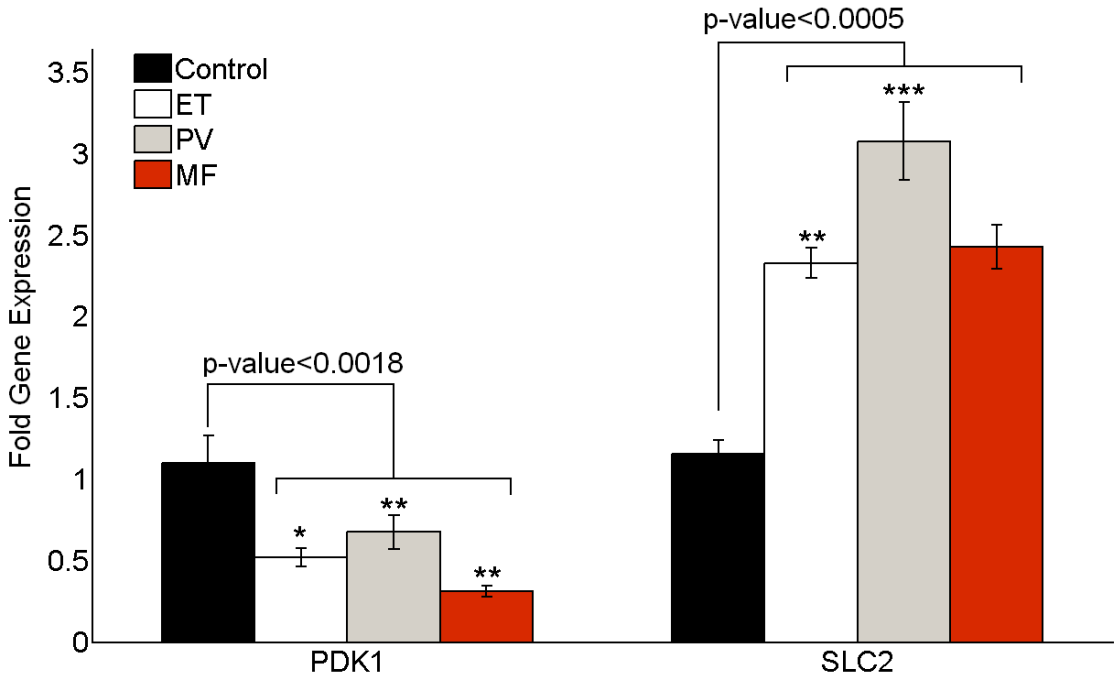


Figure 4.18: Gene expression of PDK1 and SLC2A in granulocytes from ET, PV and MF patients and controls. $*$, $**$ and $***$ correspond to p-value<0.05, p-value<0.01 and p-value<0.005, respectively, when compared to control group.

converted to Acetyl-CoA and lactate.

4.7 Discussion

The surface enhanced emission spectra measured in granulocyte has two distinctive features: (i) vibrational bands that originate from amorphous carbon structures and silver (I, II) oxide, and (ii) a broad background continuum. We showed that the broad continuum observed from granulocytes is luminescent in origin, as demonstrated by spectra acquired using different excitation wavelengths and time resolved measurements. Separation of the two contributions reveals that significant differences between MPN patients and controls arise primarily from the background continuum. For the background continuum emission, we observed a blue shift in the spectral maximum and different normalized intensity in patient groups with respect to those in the control group. The same behavior is observed in granulocytes incubated with higher 2DG concentrations. Although the origin of the background continuum emission in

SERS is still debated [8,18,26,34], it could, in part, be attributed to surface enhanced luminescence [8] of molecules close to the metallic surface (within ∼20 nm). We note that the continuum emission background is known to increase upon adsorption [8] and, under some experimental conditions, its spectral characteristics depend whether the molecule involve is an electron-donor or an electron-acceptor [34].

We believe that the background emission we observed is due to enhanced fluorescence of flavin adenine dinucleotide (FAD), which explains a lower maximum value of the background in patients' spectra as lower FAD levels have been measured in various neoplastic cells lines [35]. In correlation with the 2DG control measurements, we believe that the blue shift of the continuum background for MPN patients is due to the changes in the dielectric constant (ϵ_m) of the environment surrounding the metal-cell interface. A portion of the fluorescence emission is from a discrete electronic excited state of specific, as of yet unidentified molecules close to the metal surface. If the molecules are electron donors or acceptors with respect to the silver nanoparticles, then a charge transfer exciton (CTE) is formed at the cell/silver interface. The CTE level depends on the dielectric properties of the environment, and this in turn changes the emission spectrum from such a state. Neoplastic cells are known to have a higher ϵ_m than normal cells [36]. This change in ϵ_m, which could be due to alterations in the concentration of glycolytic metabolites and protein content [36,37], may directly change the continuum emission spectrum and thus explain the observed blue-shift. Vibrational bands at ∼1600 cm^{-1} and ∼1370 cm^{-1} indicate the formation of amorphous carbon structures; these bands allow us to differentiate between the ET group and, the PV and PMF group. We think that the discrimination is possible due to changes in concentration of the parent components used in the carbonization process such as lipids and glycolytic metabolites.

Reactive nonneoplastic subjects are a group of patients whose symptoms mimic those of MPNs but in turn are not affected by MPN. Importantly, reactive disorders are not clonal and could arise as a response to poor lung function and high altitude exposure, among others. We evaluated reactive nonneoplastic subjects as means of control. Interestingly, this group of patients seem to have lower TEI than MPN

patients, and for all ratios of intensities, they are in the extreme opposite of controls values. Further characterization of this group of patients in terms of gene expression and cellular protein contents will be required for a more complete interpretation of the data.

We observed a hypoxia-like behavior in granulocytes from MPN patients given the upregulation of SLC2A1 and the downregulation of PDK1, both glycolytic gene targets of HIF1. Our conclusions from the spectral data are supported by the expression measurements on granulocytes as a higher cellular glucose intake will be necessary for survival, explaining the shift of the background maximum, as well as the reduction of FAD, which is a consequences of the metabolic transformation [35] and explains the change in the background maximum.

4.8 Summary

In conclusion, we evaluated the diagnostic potential of normalized integrated emission and ratios of intensity pairs of the emission spectra from granulocytes of myeloproliferative disorders patients enhanced by the proximity to silver nanoparticles. We found that the intensity ratios are excellent tests for the presence of MPNs, and that some ratios were able to discriminate ET from other myeloproliferative neoplasms. We hypothesized that the spectral changes in data from MPN patients result from changes in the cellular metabolism that lead to increase glucose uptake and reduced FAD cellular content; this may explain both the continuum blue shift and the weakness of the emission background in MPN patients.

We provide additional supporting evidence for dysregulation of the expression of two genes SLC2A1 and PDK-1 regulated by HIF1 in MPNs patients suggesting that MPNs, similar to other malignancies, modify their metabolism by the Warburg effect. Therefore, our data also support the existence of the Warburg effect in MPN, which likely accounts for our SERS data.

4.9 Acknowledgements

We thank Kimberly Hickman for her invaluable help with samples and patient coordination, Dr. Nicholas Borys for assistance with time resolved measurements

and Dr. Randy Polson for assistance with EDS measurements.

4.10 References

[1] M. Lichtman, *Williams Hematology* (McGraw-Hill, Medical Pub. Division, New York, NY, 2006).

[2] P. J. Campbell and A. R. Green, "The myeloproliferative disorders," New Engl J Med **355**, 2452–2466 (2006).

[3] E. J. Baxter, L. M. Scott, and P. J. Campbell, "Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders," Lancet **365**, 1054–1061 (2005).

[4] H. Zhan and J. L. Spivak, "The diagnosis and management of polycythemia vera, essential thrombocythemia, and primary myelofibrosis in the JAK2 V617F era," Clin Adv Hematol Oncol **7**, 334–42 (2009).

[5] A. M. Vannucchi and P. Guglielmelli, "Molecular pathophysiology of Philadelphia-negative myeloproliferative disorders: beyond JAK2 and MPL mutations," Haematologica **93**, 972–6 (2008).

[6] C. Krafft, B. Dietzek, and J. Popp, "Raman and CARS microspectroscopy of cells and tissues," Analyst **134**, 1046–1057 (2009).

[7] M. Moskovits, "Surface-enhanced spectroscopy," Rev Mod Phys **57**, 783–826 (1985).

[8] M. Moskovits, "Surface-enhanced Raman spectroscopy: A brief retrospective," J Raman Spectrosc **36**, 485–496 (2005).

[9] J. Kim and C. Dang, "Cancer's molecular sweet tooth and the Warburg effect," Cancer Res **66**, 8927–8930 (2006).

[10] A. Tefferi and J. W. Vardiman, "Classification and diagnosis of myeloproliferative neoplasms: The 2008 World Health Organization criteria and point-of-care diagnostic algorithms," Leukemia **22**, 14–22 (2008).

[11] S. I. Swierczek, N. Agarwal, R. H. Nussenzveig, G. Rothstein, A. Wilson, A. Artz, and J. T. Prchal, "Hematopoiesis is not clonal in healthy elderly women," Blood **112**, 3186–3193 (2008).

[12] G. Chen and J. Prchal, "X-linked clonality testing: Interpretation and limitations," Blood **110**, 1411–1419 (2007).

[13] J. Prchal, J. F. Prchal, M. Belickova, S. Chen, Y. Guan, G. L. Gartland, and M. D. Cooper, "Clonal stability of blood cell lineages indicated by X-chromosomal transcriptional polymorphism," J Exp Med **183**, 561–567 (1996).

[14] R. H. Nussenzveig, S. I. Swierczek, J. Jelinek, A. Gaikwad, E. Liu, S. Verstovsek, J. F. Prchal, and J. T. Prchal, "Polycythemia vera is not initiated by JAK2V617F mutation," Experimental Hematology **35**, 32.e1–32.e9 (2007).

[15] B. Yeo, T. Schmid, W. Zhang, and R. Zenobi, "A strategy to prevent signal losses, analyte decomposition, and fluctuating carbon contamination bands in surface-enhanced Raman spectroscopy," Appl Spectrosc **62**, 708–713 (2008).

[16] A. Otto, "What is observed in single molecule SERS, and why?" J Raman Spectrosc **33**, 593–598 (2002).

[17] G. I. N. Waterhouse, G. A. Bowmaker, and J. B. Metson, "The thermal decomposition of silver (I, III) oxide: A combined XRD, FT-IR and Raman spectroscopic study," Phys Chem Chem Phys **3**, 3838–3845 (2001).

[18] M. L. Jacobson and K. L. Rowlen, "The role of O2 in SERS-active thin metal film photodynamics," J Phys Chem B **110**, 19 491–19 496 (2006).

[19] N. P. W. Pieczonka and R. F. Aroca, "Inherent complexities of trace detection by surface-enhanced Raman scattering," ChemPhysChem **6**, 2473–2484 (2005).

[20] X. M. Lin, Y. Cui, Y. H. Xu, B. Ren, and Z. Q. Tian, "Surface-enhanced raman spectroscopy: Substrate-related issues," Anal Bioanal Chem **394**, 1729–1745 (2009).

[21] K. L. Norrod and K. L. Rowlen, "Removal of carbonaceous contamination from SERS-active silver by self-assembly of decanethiol," Anal Chem **70**, 4218–4221 (1998).

[22] I. Pócsik, M. Veres, M. Fle, S. Tóth, and M. Koós, "Specific statistical features of surface enhanced Raman scattering (SERS) spectra of graphite," J Non-Cryst Solids **338-340**, 496–498 (2004).

[23] S. Potgieter-Vermaak, N. Maledi, N. Wagner, J. H. P. Van Heerden, R. Van Grieken, and J. H. Potgieter, "Raman spectroscopy for the analysis of coal: A review," J Raman Spectrosc **42**, 123–129 (2011).

[24] J. Zhao, H. Lui, D. I. McLean, and H. Zeng, "Automated autofluorescence background subtraction algorithm for biomedical Raman spectroscopy," Appl Spectrosc **61**, 1225–32 (2007).

[25] J. C. A. Hopkins, G. K. Radda, R. L. Veech, and K. Clarke, "Accumulation of 2-deoxy-D-glucose-6-phosphate as a measure of glucose uptake in the isolated perfused heart: A 31P NMR study," Metab Eng **6**, 36–43 (2004).

[26] N. J. Borys and J. M. Lupton, "Surface-enhanced light emission from single hot spots in tollens reaction silver nanoparticle films: Linear versus nonlinear optical excitation," J Phys Chem C **115**, 13 645–13 659 (2011).

[27] J. W. Kim, I. Tchernyshyov, G. L. Semenza, and C. V. Dang, "HIF-1-mediated expression of pyruvate dehydrogenase kinase: A metabolic switch required for cellular adaptation to hypoxia," Cell Metab **3**, 177–185 (2006).

[28] G. L. Semenza, "Signal transduction to hypoxia-inducible factor 1," Biochem Pharmacol **64**, 993–998 (2002).

[29] O. Basturk, R. Singh, E. Kaygusuz, S. Balci, N. Dursun, N. Culhaci, and N. V. Adsay, "GLUT-1 expression in pancreatic neoplasia: Implications in pathogenesis, diagnosis, and prognosis," Pancreas **40**, 187–192 (2011).

[30] H. Ito, M. Duxbury, M. J. Zinner, S. W. Ashley, and E. E. Whang, "Glucose transporter-1 gene expression is associated with pancreatic cancer invasiveness and MMP-2 activity," Surgery **136**, 548–556 (2004).

[31] T. Kawamura, T. Kusakabe, T. Sugino, K. Watanabe, T. Fukuda, A. Nashimoto, K. Honma, and T. Suzuki, "Expression of glucose transporter-1 in human gastric carcinoma: Association with tumor aggressiveness, metastasis, and patient survival," Cancer **92**, 634–641 (2001).

[32] M. Younes, R. W. Brown, M. Stephenson, M. Gondo, and P. T. Cagle, "Overexpression of Glut1 and Glut3 in stage I nonsmall cell lung carcinoma is associated with poor survival," Cancer **80**, 1046–1051 (1997).

[33] M. Younes, D. Juarez, L. V. Lechago, and S. P. Lerner, "Glut 1 expression in transitional cell carcinoma of the urinary bladder is associated with poor patient survival," Anticancer Res **21**, 575–578 (2001).

[34] S. Mahajan, R. M. Cole, J. D. Speed, S. H. Pelfrey, A. E. Russell, P. N. Bartlett, S. M. Barnett, and J. J. Baumberg, "Understanding the surface-enhanced Raman spectroscopy "background"," J Phys Chem C **114**, 7242–7250 (2010).

[35] J. H. Ostrander, C. M. McMahon, S. Lem, S. Millon, J. Brown, V. L. Seewaldt, and N. Ramanujam, "Optical redox ratio differentiates breast cancer cell lines based on estrogen receptor status," Cancer Res **70**, 4759–4766 (2010).

[36] X. J. Liang, A. Q. Liu, C. S. Lim, T. C. Ayi, and P. H. Yap, "Determining refractive index of single living cell using an integrated microchip," Sensors Actuat A-Phys **133**, 349–354 (2007).

[37] R. Thar and M. Khl, "Propagation of electromagnetic radiation in mitochondria?" J Theor Biology **230**, 261–270 (2004).

CHAPTER 5

SURFACE ENHANCED RAMAN SPECTROSCOPY: OTHER APPLICATIONS

Label-free surface enhanced raman spectroscopy for the detection of proteins in solutions is highly desirable as it would simplify sample preparation and data analysis for medical screening and diagnostic applications. In this chapter, we describe our studies related to the application of SERS for the detection of hemoglobin (Hb) variants and the prediction of subjects' response to the seasonal influenza vaccine.

5.1 Detection of hemoglobin variants

As described in section 2.5, the detection of Hb variants is not always possible by the standard electrophoretic or chromatographic methods. Hb variants are molecules that resemble the normal hemoglobin except for one amino acid substitution. Thus, differences in mass can be small enough to elude mass spectrometry. Hb variants are structurally different when compared to normal hemoglobin, and to one another. So, we propose SERS as an alternative for the detection of Hb variants that has the potential to differentiate them in a quantitative fashion.

5.1.1 Methods

Whole blood samples were received from ARUP laboratories (Salt Lake City, UT) and hemolysate was obtained the same day using the protocol described below. In addition, standard human hemoglobin (Sigma Aldrich, St Louis, MO) was purchased as a reference.

5.1.1.1 Hemolysate protocol

2 ml of whole blood were diluted in 6 ml of PBS at room temperature. The solution was centrifuged at 600 g for 10 min, and the supernatant was discarded. Steps 1 and 2 were repeated at least twice. In between centrifugation steps, blood was kept on ice. Pellet was diluted in autoclaved distilled water at 4 ^{o}C to fill a 15 ml centrifuge tube up to 10 ml. The dilution stood for 15 min. Samples were centrifuged at 2000 g for 15 min. The top 5 to 6 ml were collected and transferred to a new tube, while the rest was discarded.

The confocal microspectroscopy configuration described in Section 3.1.1 was used for SERS spectra acquisition using a microscope objective of 60X (NA 0.80). Silver nanoparticles deposited on glass (i.e., substrates) were prepared by the Tollens reaction as explained in Section 3.1.2 with a 120 seconds reaction time. The substrates were stored for a time period shorter than one day under a nitrogen atmosphere from the time of preparation until their use. A drop of hemolysate was placed onto the silver nanoparticles substrate, and a coverslip was place on top, carefully to avoiding air bubbles. To acquire the SERS spectra, 5 to 6 scans of 25 μm by 25 μm (20 by 20 pixels) per sample were done, collecting 400 spectra per scan. Not all of the acquired spectra were surfaced-enhanced. Thus, a selection algorithm was implemented based on the spectral flatness, defined as the ratio of the geometrical mean to the arithmetical mean of the spectrum (eq. 5.1) [1].

$$flatness = \frac{\sqrt[n]{\prod_n x_n}}{\frac{\sum_n x_n}{n}} \tag{5.1}$$

where n is the the number of points per spectrum. Using this criterion, only between 0.1 % and 2 % of the total number of spectra per sample were selected. Selected SERS spectra were normalized to the total integrated emission and averaged, such that each sample was represented by one spectrum during principal component analysis (PCA).

PCA is explained in detail in section 3.2.3. Briefly, PCA works as a linear transformation from the space of variables to a new space that represents the data in a more meaningful way. This representation has its origin in the center of gravity of the data

points and has a new basis of which the axes are called *principal components*(PCs). PCs lie along the directions of maximum variability in the data. PC1 is the one that describes most of the variance, followed by PC2, PC3 and so forth. Each data point in the PCs space is thus represented by a new set of coordinates called scores. Plots of scores provide relevant information on hidden patterns in the data, such as correlation or clustering for example. The loadings plot of a PC maps the contributions of each variable to the PC.

Classification of various Hb variants was explored using support vector machine (SVM) classification [2]. Briefly, SVM involves the construction of a hyperplane[1] in such way that the distance from the closest samples to the hyperplane from two different classes is maximized; a 2-dimensional (2D) example is shown in Figure 5.1. Here, SVM is used to show that classification of hemoglobin variants based on the SERS spectra is possible. The results from the SVM are given as a *'confusion matrix'* which is explained in Table 5.1. The miss rate defined in equation 5.2 and the positive predictive value (PPV) and negative predictive value (NPV), are defined in equations 5.3 and 5.4, respectively. Both PCA and SVM computation were done with the *The Unscrambler X* software [3].

$$Miss\ Rate = \frac{Number\ of\ miss\ assigned\ cases}{Total\ number\ of\ cases} \times 100 \tag{5.2}$$

$$PPV = \frac{Number\ of\ true\ positive\ cases}{Total\ number\ of\ positive\ cases} \times 100 \tag{5.3}$$

$$NPV = \frac{Number\ of\ true\ negative\ cases}{Total\ number\ of\ negative\ cases} \times 100 \tag{5.4}$$

Low PVV or NPV indicate a high number of cases being incorrectly assigned as positive or negative cases, respectively. In the case of low PPV, follow up tests are necessary to confirm a positive outcome, whereas for low NPV, additional tests are

[1]The hyperplane of a n-dimensional space is a subspace with n-1 dimensions that divides the space in two.

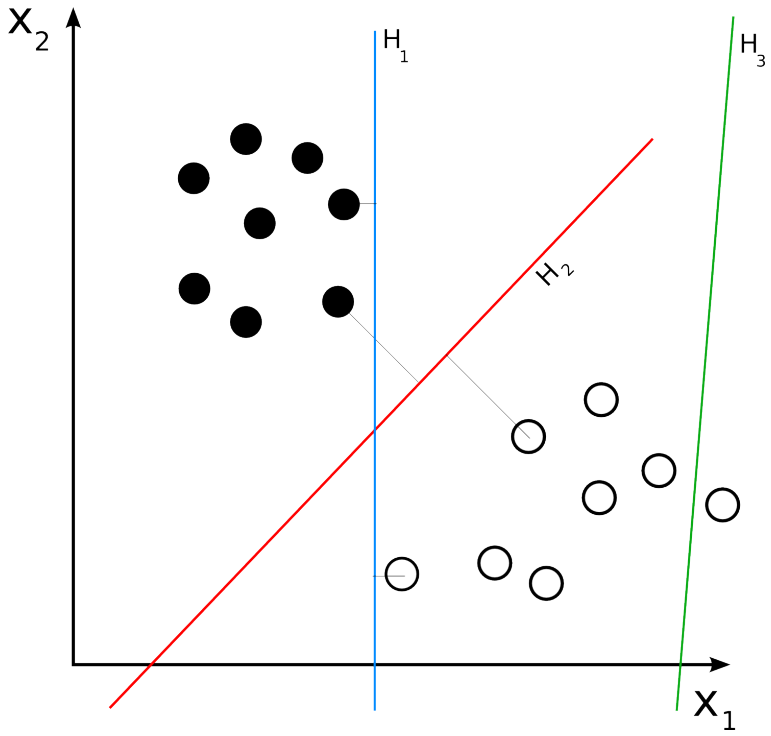


Figure 5.1: The principle of support vector machine (SVM) classification techniques is exemplified here. For a 2D space the hyperplane is a straight line, but only one hyperplane, H_2 in this case, maximizes the distance from it to the closest samples in both groups.

Table 5.1: Confusion matrix for two hemoglobin variants

		Actual	
Predicted		Number of true Hb E cases	Number of False Hb S cases
		Number of False Hb E cases	Number of True Hb S cases

required to confirm negative result. Screening test refers to tests that generally have either high PPV or NPV, but not both.

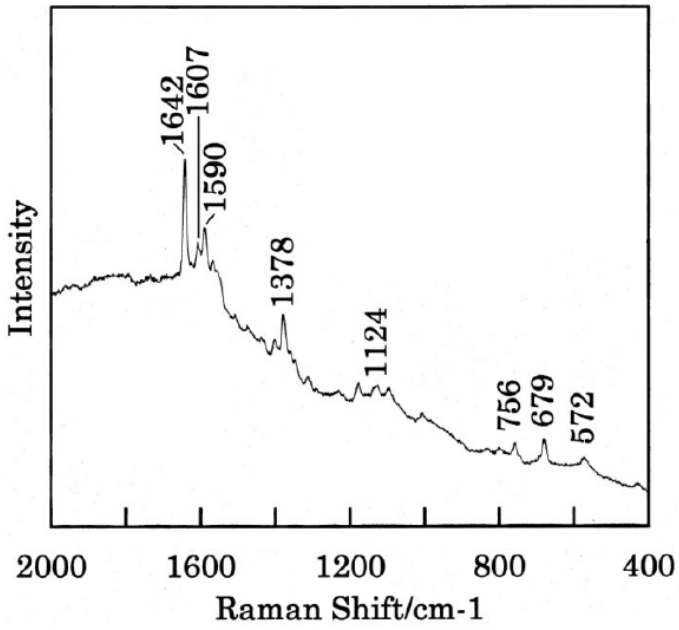
5.1.1.2 Silver nanoparticles colloid

In addition to the measurements performed using the Tollens substrate, measurements on standard, normal human hemoglobin were done. The silver colloid substrate was prepared according to published protocol [4]. Briefly, 0.1 mM of aqueous silver nitrate was heated to 98 ^{o}C, at which point 2 ml of 1 % Trisodium Citrate solution were added. The mixture was kept at constant temperature of 98 ^{o}C for 60 min, and special care was taken to prevent the mixture from boiling. In the case of colloidal nanoparticles, aggregation agents are required to maintain a short distance between the analyte and the metal nanoparticles. We used the aggregation agent reported by Han *et al.* [5]. The aggregation agent is 0.1 M aqueous solution of sodium sulfate with pH of 3 which was adjusted using sulfuric acid.

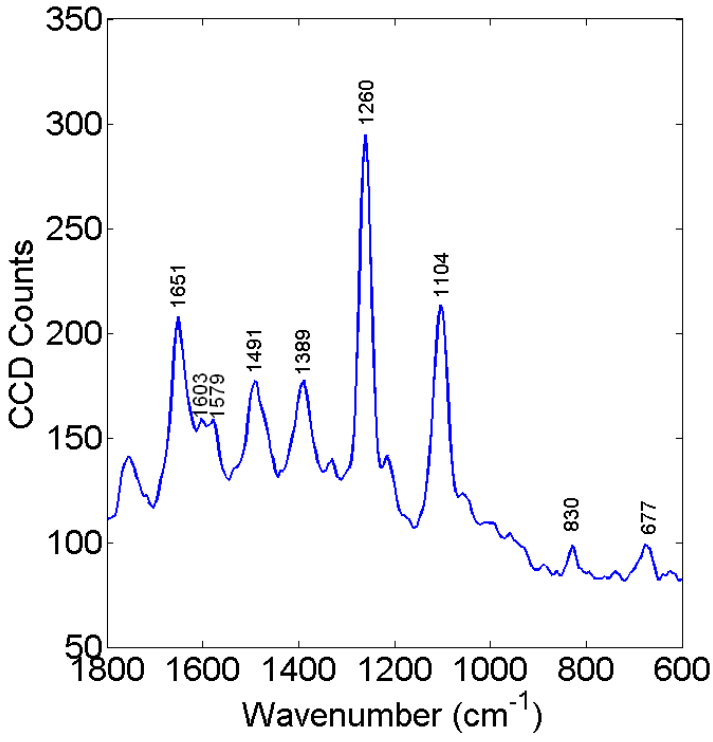
A solution of hemoglobin in water was prepared and mixed in a 1:10 ratio with the aggregation agent. The Hb/aggregation agent was mixed with a 1:5 volume ratio with the silver nanoparticles colloid. The SERS measurements were done immediately following the preparation procedure.

5.1.2 Results and discussion

Figure 5.2B shows the SERS spectrum of hemoglobin excited with a 488 nm argon laser and a power of 0.5 mW. For comparison the Raman spectrum of hemoglobin excited with 514.5 nm laser is also shown (Figure 5.2A). Notice that the vibrational bands in Figure 5.2B are not equally enhanced; for example, the SERS bands at 830



A



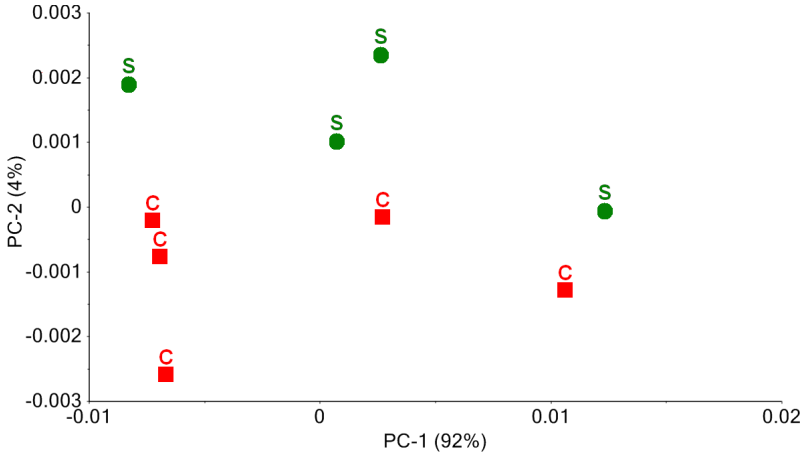
B

Figure 5.2: Spectra of hemoglobin. **A:** Normal Raman spectrum excited at $\lambda_{exc} = 514.5nm$. Reprinted with permission from [6]. Copyright 2001 by SPIE. **B:** Measured SERS of hemoglobin excited at $\lambda_{exc} = 488nm$ on Tollens substrate.

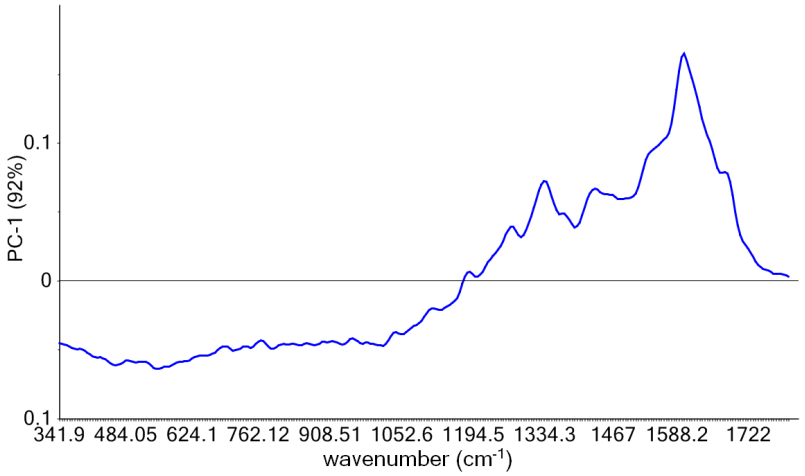
cm^{-1} and 1491 cm^{-1} are barely visible in the regular Raman spectrum. This is due to adsorption of the hemoglobin onto the silver, which in turn changes the Raman cross-section of the molecule, and the proximity to the nanoparticles selectively enhances some of the vibrational bands [7].

Hb variants are structurally different. The differences arise from a single amino acid substitution in the protein. Since the difference is structural, SERS is expected to be able to discriminate among various Hb variants. This could successively help the diagnosis of those variants which elude other techniques, as explained before. To prove the ability of SERS to discriminate Hb variants, samples from patients affected with Hb E [8], Hb S [9], Hb C [10] and traits FE and FS were analyzed as follows. The results are shown in Figures 5.3 through 5.5, where the scores plots and the corresponding loading plots are shown. Figure 5.3A shows that *Hb variants S and C can be discriminated* based on PC1 and PC2 scores, which together explain 98% of the variability in the data set. Figure 5.3B and 5.3C show that a large part of the variability in the case of Hb S and Hb C, is explained by the spectral features in the 1400-1800 cm^{-1} region and peaks at 1603 cm^{-1}. SVM analysis was done on the data sets shown in Figures 5.3A, 5.4 and 5.5. The results are shown in Table 5.2. For the case of hemoglobin E and traits FS and FE, even though a visible separation is clear in Figure 5.5B, SVM results were poor and have a miss rate of 71 %. On the other hand, when the only FE sample was taken out of the data set, the miss rate improved to 17 %.

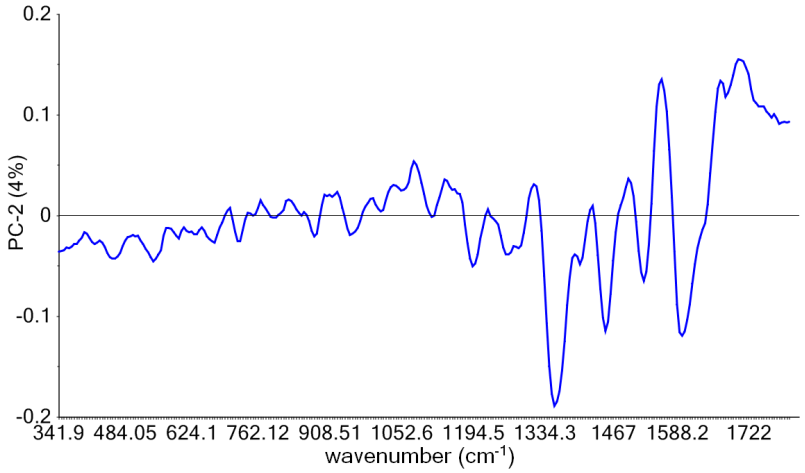
These results provide evidence that SERS can distinguish between different variants of hemoglobin. In addition, we also explored silver nanoparticle colloid as means for Raman scattering enhancement. Figure 5.6A shows SERS spectra from standard human hemoglobin measured using a 488 nm argon laser at a power of 0.5 mW. Characteristic spectral features of hemoglobin are shown in the region between 1300 and 1700 wavenumbers. The features in the region 800-1000 cm^{-1} arise from the aggregation agent and the silver colloid, whereas the broad band at $\sim$500 cm^{-1} is attributed to silver oxides [11] as they are observed from the colloid/aggregation agent mixture in the absence of an analyte. For comparison, Figure 5.6B shows the



A



B



C

Figure 5.3: PCA for Hb variant S (green) and Hb variant C (red). **A:** Scores plot PC1 vs. PC2, **B:** loading plot spectrum for PC1 and **C:** loading plot spectrum for PC2

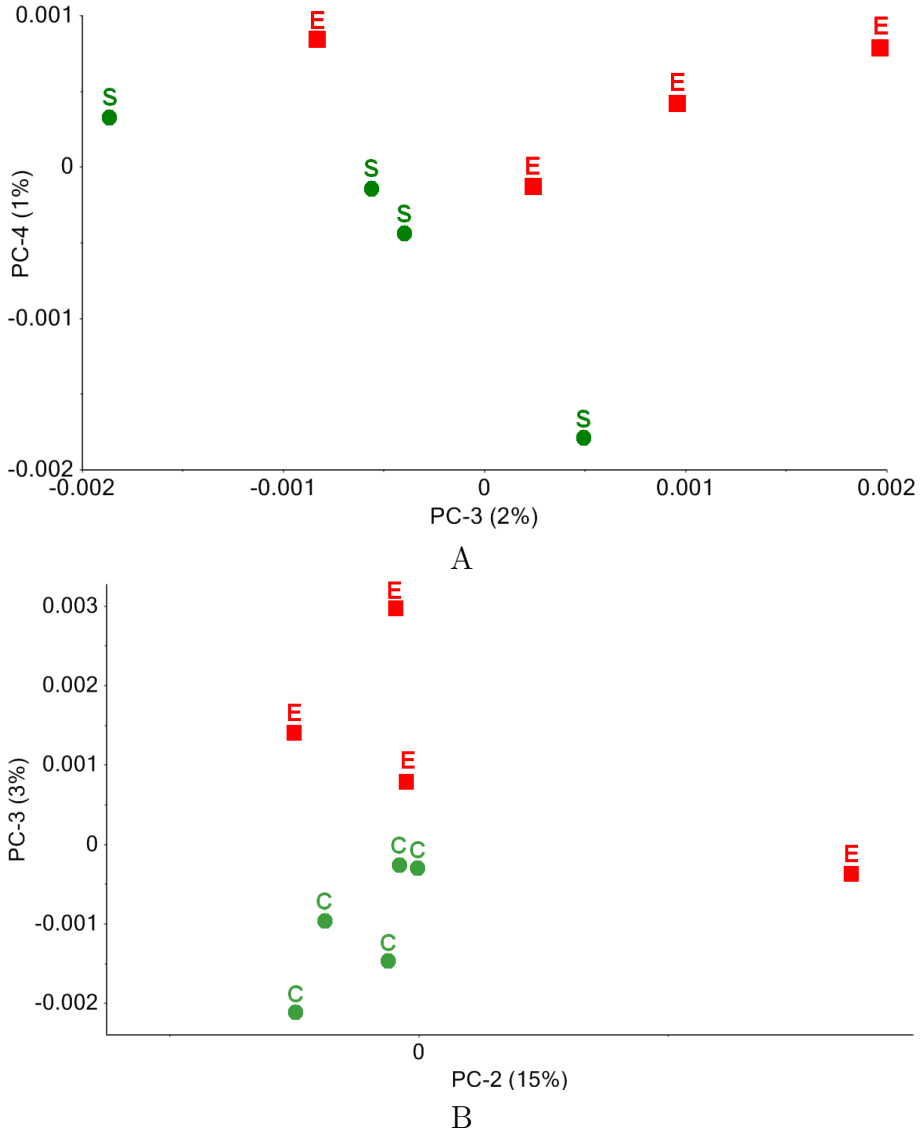


Figure 5.4: Scores plots of **A:** PC4 vs. PC3 for Hb variant S (green) and Hb variant E (red) and **B:** PC3 vs. PC2 for Hb variant C(green) and Hb variant E (red).

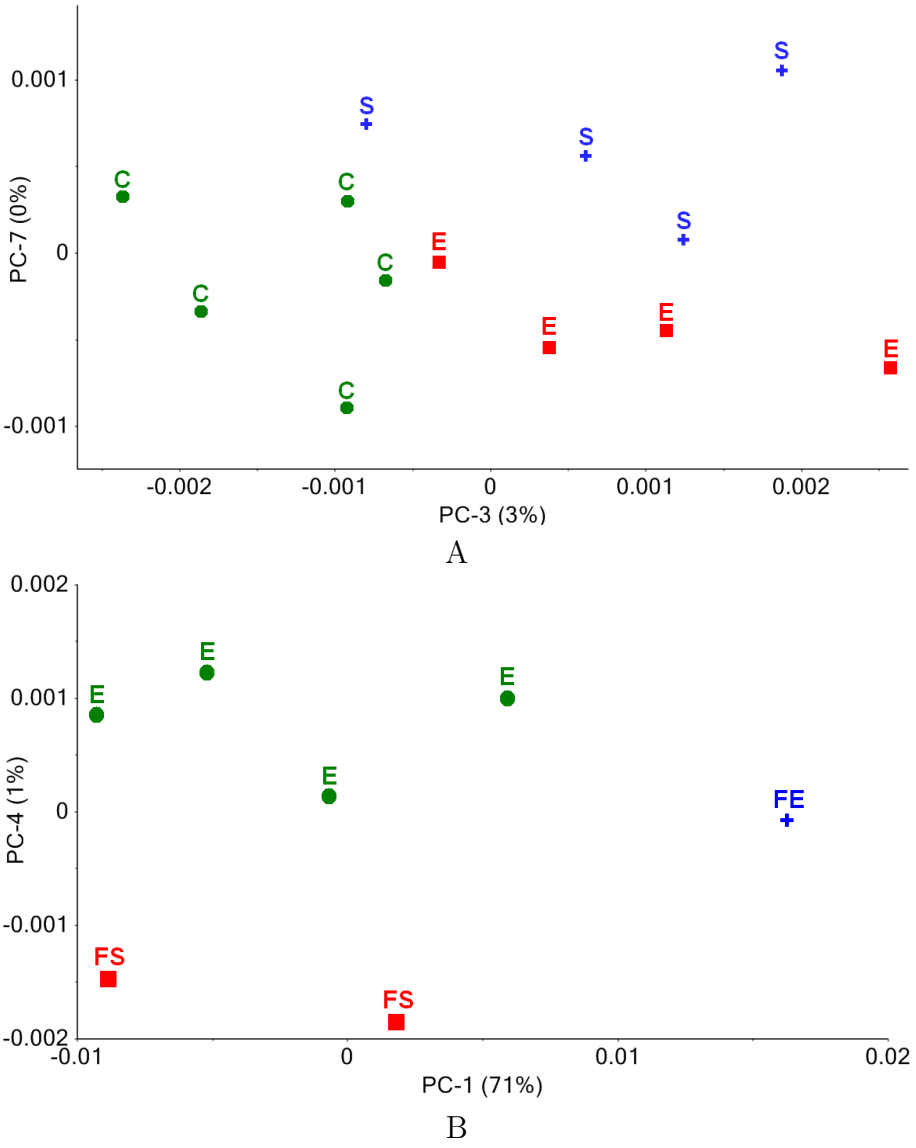


Figure 5.5: Scores plots of **A:** PC7 vs. PC3 for Hb variants C(green), Hb E(red) and Hb S(blue) and **B:** PC7 vs. PC3 for Hb E(green) and traits FS(red) and FE(blue).

Table 5.2: SVM results for hemoglobin variants. The 'miss rate' is defined in equation 5.2.

Data Set	No. False Cases	Total Cases	Miss Rate
Hb E, Hb C	1	9	11%
Hb E, Hb S	1	8	12.5%
Hb S, Hb C	1	9	11%
Hb E, Hb C, Hb S	3	13	23%

Raman scattering spectra from crystalized hemoglobin (acquired by Xu *et al*), and hemoglobin/silver aggregates in solution excited with 514.5 nm laser wavelength at 1 mW and 1 μW of excitation power, respectively. It is evident that SERS is not observed in this case, possibly because single silver nanoparticles are not as good for enhancement of the Hb Raman spectrum, and/or the concentration of Hb is too high, thus recovering bulk Hb Raman [12]. Very low concentrations of hemoglobin were tried, and similar spectra to the one shown in Figure 5.6A were obtained. Failed SERS measurements have been also reported in ref. [13]. Xu *et al.* reported single hemoglobin molecule SERS from *deposited* mixtures of silver nanoparticles colloid with very low concentrations of hemoglobin [12]. These results suggest that SERS measurements from hemoglobin using colloidal silver nanoparticles in solution are not feasible.

5.2 Seasonal influenza response[3]

Chronic heart failure (CHF) is considered a high-risk condition for influenza illness. Vaccination against influenza infection results in decreased cardiac related hospital admissions, a reduction in acute heart failure exacerbations, and is known to decrease all cause mortality [14]. Older adults, and persons of all ages with cardiac disease or other co-morbidities and treatments that render them immune-compromised, are known to be at greater risk for influenza infection despite being vaccinated due to reduced antibody and cell mediated responses to vaccines [15, 16]. Studying the

[3]This project was done in collaboration with Professor Orly Vardeny from the School of Pharmacy at the University of Wisconsin-Madison.

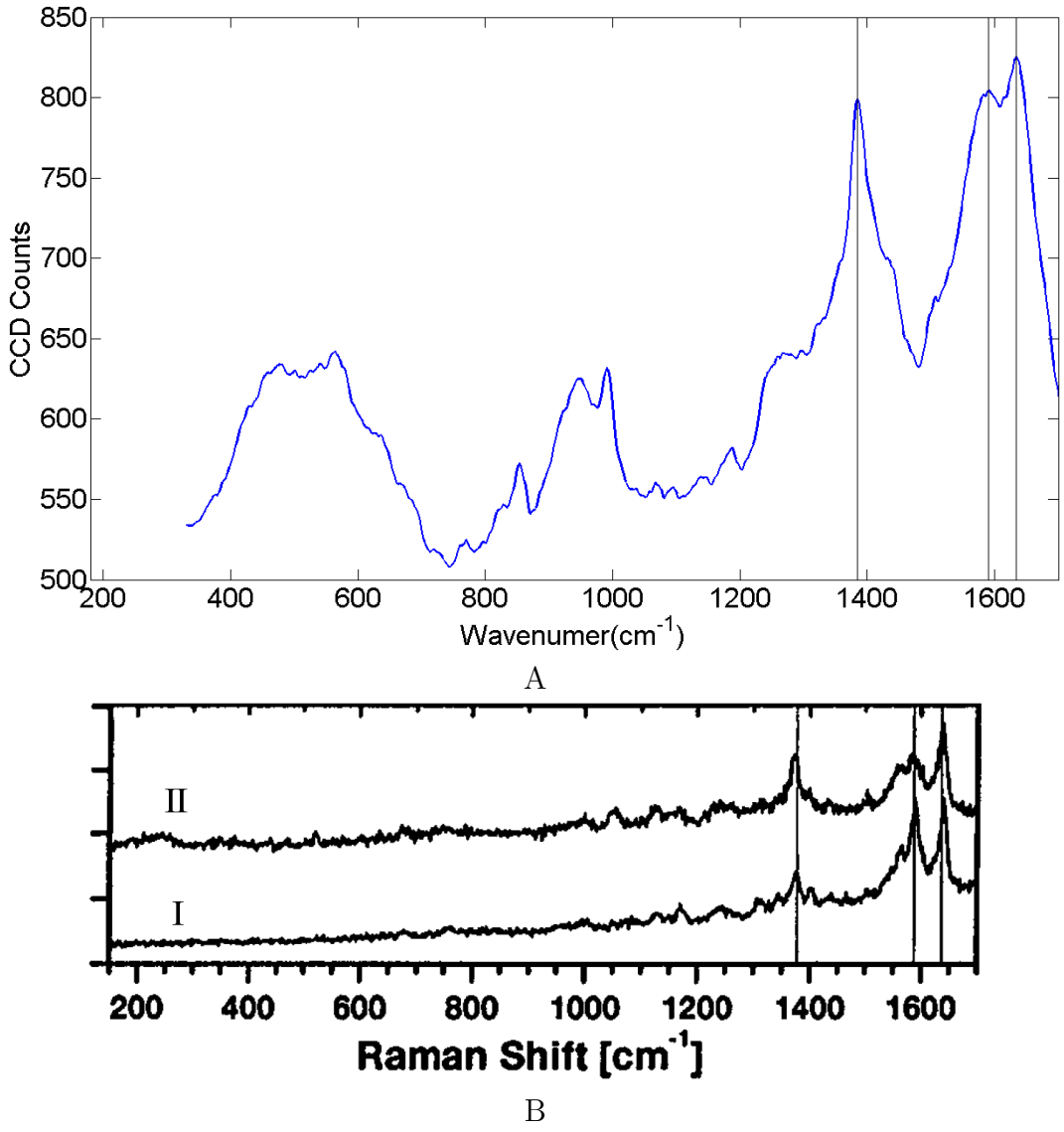


Figure 5.6: SERS spectra of hemoglobin. **A:**Raman spectrum from a mixture of hemoglobin and silver nanoparticles colloidal acquired using 488nm excitation wavelength at 0.5mW power. **B:** Spectra from crystalized hemoglobin (I) and hemoglobin/silver aggregates in solution (II) with 514.5nm excitation wavelength and 1mW (I) and $1\mu W$ (II) excitation power, respectively. Reprinted with permission from [12]. Copyright 1999 by the American Physical Society.

response to the influenza vaccine within this population is of medical relevance since the results may contribute to the future planning in vaccination strategies. The usage of SERS, Raman scattering and autofluorescence spectroscopies may contribute to the advancement of new research and diagnosis technologies.

Immunoglobulins (Igs), i.e., antibodies, are proteins found in the blood and body fluids that are part of the immune system and help recognize foreign objects, such as viruses and bacteria. Antibodies are secreted by B-cells in the blood, which can produce as many as 10^{14} different antibodies. In contrast, an antigen is a molecule that distinguishes the foreign object. Thus, when the foreign object enters the body it is recognized by antibodies, consequently triggering the immune system response. The antibody recognizes the antigen through a highly specific binding process called '*induced fit*'. Specific antibodies are used as indicators of bacterial or viral infections, as well as a measure for response to vaccines. Enzyme-linked immunosorbent assay (ELISA) is a widely used method to detect antibodies. Additional methods used are chemiluminescence assay, enzymed-linked fluorescence assay and indirect immunofluorescent assay. All of these methods provide high sensitivity and specificity. However, these methods are not sufficiently fast because of very specific and multiple sample preparation steps [17, 18].

Recently, it has been shown that SERS allows detection of antibodies in human colostrum [19] and human serum [20], as well as antibody-antigen interaction [21]. Based on these findings, heart failure patients are expected to show different spectra than age-matched controls when measured on blood serum, due to differences in concentrations of antibodies.

In the present study SERS measurements are demonstrated to correlate with a positive/negative response to the seasonal influenza vaccine. Measurements were done on samples collected before and after influenza vaccine administration and provided by Dr. Orly Vardeny from the University of Wisconsin-Madison.

5.2.1 Methods

Serum samples were obtained from CHF patients and healthy subjects before and after seasonal influenza vaccination. Sample collection was done at the University of

Wisconsin-Madison, and the project underwent review by IRBs at both institutions. Before and after vaccination serum samples were analyzed by ELISA method, and determination of a positive or negative response based on differences in antibody concentrations pre- and postvaccination was made [22]. Confocal microspectroscopy was used to acquire emission spectra under the same conditions as described in section 5.1.1. A drop of serum diluted with a ratio of 1:1 in ultra pure water was placed onto the Tollens substrate, and a coverslip placed on top. The spectra selection criteria were given by spectral flatness, eq. 5.1, as described above. After selection, the average spectra were calculated per subject, before and after vaccination. Average SERS spectra were analyzed using PCA and SVM (sections 3.2.3 and 5.1.1).

5.2.2 Results and discussion

The ability of SERS to differentiate among subjects who responded to the vaccines from those who did not was tested. SVM analysis was done on the the first 7 principal components, which together explained 100 % of the variability within the data. PCA and subsequent binary classification among responders and nonresponders was done within the subpopulations as described in Table 5.3.

For the first group, that is all subjects before vaccination, false positives were not observed. On the other hand, NPV was 75%. This result implies that *SERS on serum can distinguish whether a subject will respond to the vaccine*. PPV and NPV had the same values after vaccination as they did before vaccination, i.e., 100 % and 75 %, respectively. When SERS spectra were analyzed exclusively from CHF patients, PPV and NPV were 100% and 71.4%, respectively, for both pre- and postvaccination. The

Table 5.3: Subpopulations and groups defined therein used for PCA and subsequent SMV analysis of SERS spectra of human serum

Analysis	Subpopulation
1	All subjects before vaccination
2	All subjects after vaccination
3	CHF Patients before vaccination
4	CHF Patients after vaccination

distinction between responders and nonresponders is depicted in Figure 5.7, which shows the scores plot of PC2 vs PC1 of CHF patients' spectra after vaccination.

These results open the possibility for a screening test that will allow *prediction of response to seasonal influenza vaccines*, which in turn could help to improve the care management in CHF patients. In addition, using a different approach for data analysis such as regression methods, SERS has the potential of measuring specific antibodies, such as in this case, those associated with the influenza strains included in each year's vaccine. Similar experiments have shown that it is possible to measure the concentration of various chemicals in blood serum and urine by normal Raman scattering [23]; to detect clenbuterol using SERS in conjunction with functionalized nanoparticles [24]; to detect of buserelin with a technique that combines SERS and electrophoresis techniques and allows a limit of detection as low as picomols per liter in highly complex samples, such as body fluids [25]; and to classify blood serum from subjects that are either hepatitis C positive or negative based on SERS.

SERS measurements on hemoglobin variants and pre- and postseasonal influenza vaccination blood serum demonstrate the ability of this technique in combination with multivariate statistical analysis and classification algorithms to separate and distinguish blood samples with a variety of different characteristics.

5.3 References

[1] J. D. Johnston, "Transform coding of audio signals using perceptual noise criteria," IEEE J Sel Area Comm **6**, 314–323 (1988).

[2] A. I. Belousov, S. A. Verzakov, and J. von Frese, "Applicational aspects of support vector machines," J Chemometr **16**, 482–489 (2002).

[3] C. S. AS, *The Unscrambler X* (Oslo, Norway, 2009-2010), build version: 0.0.0.42 edn.

[4] M. Cyrankiewicz, T. Wybranowski, and S. Kruszewski, "Study of SERS efficiency of metallic colloidal systems," J Phys Conf Ser **79** (2007).

[5] X. X. Han, G. G. Huang, B. Zhao, and Y. Ozaki, "Label-free highly sensitive detection of proteins in aqueous solutions using surface-enhanced raman scattering," Anal Chem **81**, 3329–3333 (2009).

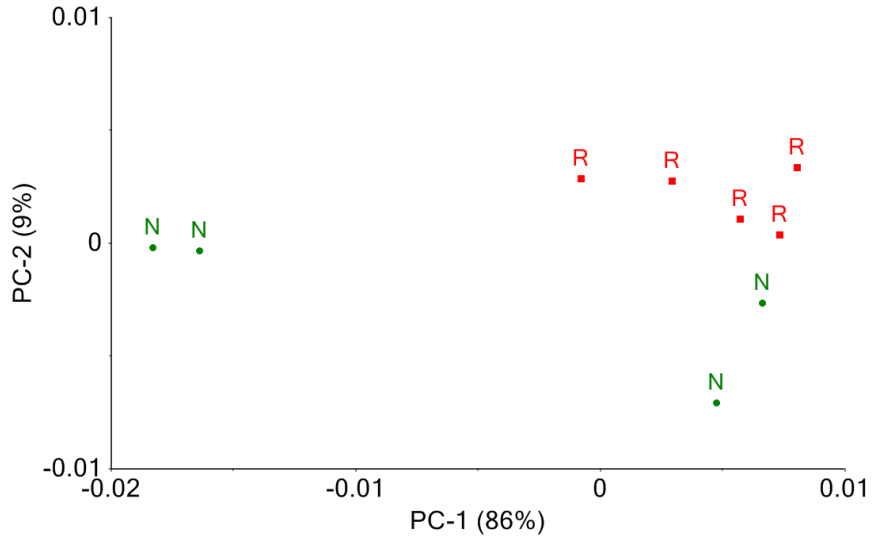


Figure 5.7: Scores plot of PC2 vs PC1 for SERS spectra from serum of CHF patients spectra after vaccination. R=Responder and N=nonresponder.

[6] H. Sato, H. Chiba, H. Tashiro, and Y. Ozaki, "Excitation wavelength-dependent changes in Raman spectra of whole blood and hemoglobin: Comparison of the spectra with 514.5-, 720-, and 1064-nm excitation," J Biomed Opt **6**, 366–370 (2001).

[7] M. Moskovits, "Surface-enhanced Raman spectroscopy: A brief retrospective," J Raman Spectrosc **36**, 485–496 (2005).

[8] J. A. Hunt and V. M. Ingram, "Abnormal human haemoglobins VI. The chemical difference between haemoglobins A and E," Biochim Biophys Acta **49**, 520–536 (1961).

[9] V. M. Ingram, "Abnormal human haemoglobins. III the chemical difference between normal and sickle cell haemoglobins," Biochim Biophys Acta **36**, 402–411 (1959).

[10] J. A. Hunt and V. M. Ingram, "Abnormal human haemoglobins: IV. The chemical difference between normal human haemoglobin and haemoglobin C," Biochim Biophys Acta **42**, 409–421 (1960).

[11] M. L. Jacobson and K. L. Rowlen, "The role of O2 in SERS-active thin metal film photodynamics," J Phys Chem B **110**, 19 491–19 496 (2006).

[12] H. Xu, E. J. Bjerneld, M. Kll, and L. Brjesson, "Spectroscopy of single hemoglobin molecules by surface enhanced Raman scattering," PRL **83**, 4357–4360 (1999).

[13] Y. Kang, M. Si, R. Liu, and S. Qiao, "Surface-enhanced Raman scattering (SERS) spectra of hemoglobin on nano silver film prepared by electrolysis method," J Raman Spectrosc **41**, 614–617 (2010).

[14] K. L. Nichol, J. Nordin, J. Mullooly, R. Lask, K. Fillbrandt, and M. Iwane, "Influenza vaccination and reduction in hospitalizations for cardiac disease and stroke among the elderly," New Engl J Med **348**, 1322–1332 (2003).

[15] A. H. Ahmed, K. G. Nicholson, J. S. Nguyen-Van Tam, and J. C. G. Pearson, "Effectiveness of influenza vaccine in reducing hospital admissions during the 1989-90 epidemic," Epidemiol Infec **118**, 27–33 (1997).

[16] E. Hak, E. Buskens, G. A. Van Essen, D. H. De Bakker, D. E. Grobbee, M. A. J. B. Tacken, B. A. Van Hout, and T. J. M. Verheij, "Clinical effectiveness of influenza vaccination in persons younger than 65 years with high-risk medical conditions: The PRISMA study," Arch Inter Med **165**, 274–280 (2005).

[17] J. Duarte, M. T. Pacheco, A. B. Villaverde, R. Z. Machado, R. A. Zangaro, and L. Silveira Jr, "Near-infrared Raman spectroscopy to detect anti-Toxoplasma gondii antibody in blood sera of domestic cats: Quantitative analysis based on partial least-squares multivariate statistics," J Biomed Opt **15**, 047 002 (2010).

[18] P. J. Lambert, A. G. Whitman, O. F. Dyson, and S. M. Akula, "Raman spectroscopy: The gateway into tomorrow's virology," Virol J **3** (2006).

[19] C. Araujo-Andrade, J. L. Pichardo-Molina, G. Barbosa-Sabanero, C. Frausto-Reyes, and A. Torres-López, "Detection of the presence of antibodies against Toxoplasma gondii in human colostrum by Raman spectroscopy and principal component analysis," J Biomed Opt **12** (2007).

[20] J. Saade, M. T. T. Pacheco, M. R. Rodrigues, and L. Silveira Jr, "Identification of hepatitis C in human blood serum by near-infrared Raman spectroscopy," Spectroscopy **22**, 387–395 (2008).

[21] K. Fujiwara, H. Watarai, H. Itoh, E. Nakahama, and N. Ogawa, "Measurement of antibody binding to protein immobilized on gold nanoparticles by localized surface plasmon spectroscopy," Anal Bioanal Chem **386**, 639–644 (2006).

[22] O. Vardeny, J. J. M. Moran, N. K. Sweitzer, M. R. Johnson, and M. S. Hayney, "Decreased T-cell responses to influenza vaccination in patients with heart failure," Pharmacotherapy **30**, 10–16 (2010).

[23] D. Qi and A. J. Berger, "Chemical concentration measurement in blood serum and urine samples using liquid-core optical fiber Raman spectroscopy," Appl Optics **46**, 1726–1734 (2007).

[24] G. Zhu, Y. Hu, J. Gao, and L. Zhong, "Highly sensitive detection of clenbuterol using competitive surface-enhanced Raman scattering immunoassay," Anal Chim Acta **697**, 61–66 (2011).

[25] V. Ranc, A. Stanova, J. Marak, V. Maier, J. Sevcík, and D. Kaniansky, "Preparative isotachophoresis with surface enhanced Raman scattering as a promising tool for clinical samples analysis," J Chromatogr A **1218**, 205–210 (2011).

CHAPTER 6

SUMMARY AND FUTURE DIRECTIONS

6.1 Summary

Autofluorescence, Raman scattering and SERS, among other optical techniques, have been proven to be promising tools for medical diagnosis and biological sensing [1–9], providing high sensitivity and specificity as well as fast and relatively simple implementation strategies. This work encompasses the development of medical diagnostic tools based on fluorescence and SERS that have the potential of greatly benefiting the care management of people affected with myeloproliferative disorders, hemoglobinopathies and chronic heart failure.

6.1.1 Myeloproliferative neoplasms

MPNs are hematological malignancies due to clonal proliferation that arises from somatic mutation of a single stem cell. These diseases result in clonal granulocytes and other circulating blood cells that are often difficult and expensive to differentiate from reactive cells as well as from congenital disorders caused by germline mutations. We presented an evaluation of SERS spectroscopy for differentiation of MPN granulocytes from controls. We found that several spectral ratios of the background fluorescence intensity that accompanies the SERS spectra of granulocytes serve as excellent markers for the presence of MPNs. In addition, the total emission intensity obtained from the normalized area under the SERS spectrum is also a good marker for MPNs, although 'specificity' and 'sensitivity' are lower than those obtained using the intensity ratios. We found expression dysregulation of two HIF-regulated genes, namely SLC2A1 and PDK1, which correlate with our results obtained by SERS assay in MPN patients and support the presence of the Warburg effect in clonal granulocytes.

We hypothesized that SERS measures metabolic change in granulocytes through

two possible mechanisms: (i) Changes in dielectric properties of the environment surrounding the silver-cell interface; and (ii) changes in flavin adenine dinucleotide concentrations, which in turn changes the relative contribution of the autofluorescence to the emission spectrum. These hypotheses are supported by SERS measurements of 2-deoxy-D-glucose incubated granulocytes where the emission spectra show a similar behavior as to that observed in the SERS spectra of controls and patients.

6.1.2 Hemoglobinopathies and chronic heart failure

Label-free SERS for the detection of proteins in solutions is highly desirable as it may simplify sample preparation and data analysis for medical screening and diagnostic applications. We presented applications of SERS towards the detection of hemoglobin variants, and screening of CHF patient who might not respond to the seasonal influenza vaccine.

Hemoglobin variants are molecules that resemble the normal hemoglobin except for one amino acid substitution. Common methods used for the detection of Hb variants are electrophoresis and chromatography and recently, mass spectroscopy [10]. Some Hb variants elude all methods, making proper diagnosis rather difficult. We proposed SERS as an alternative for Hb Variant detection and we obtained excellent results. We were able to discriminate among Hb variants S, C and E as well as traits FS and FE using classification techniques such as supporting vector machine.

Older adults, and persons of all ages with cardiac disease or other co-morbidities and treatments that render them immune-compromised, are known to be at greater risk for influenza infection despite being vaccinated due to reduced antibody and cell mediated responses to vaccines [11,12]. Studying the response to the influenza vaccine within this population is of medical relevance since the results may contribute to the future planning in vaccination strategies. We proposed SERS as a technique for the screening of responsive and nonresponsive patients to the seasonal influenza vaccine. Based on multivariate analysis of SERS spectra, we were able to distinguish those subjects who did not respond to vaccine, even before the vaccination. This technique has the potential to inform health care practitioners of people who are at risk of not responding to the influenza vaccine and therefore take the necessary precautions.

Moreover, SERS has great quantitative capabilities as demonstrated by work done on body fluids by other research groups [7–9, 13–15]. Refinement of the strategy used for this work could lead to detection, as well as quantification, of specific antibodies.

6.2 Future directions

As for every research project many questions emerged as we learned more about the techniques and their capabilities. Here we propose some projects that could follow up the work described in this thesis.

In general, all the projects would benefit from larger sample statistics, especially the tests that involved hemoglobin variant detection and CHF patient screening. A larger sample size would not just help to improve the confidence in the results, but would allow the construction of regression models that in turn could be used for quantitative sensing of hemoglobin and/or antibody concentration. A comprehensive survey of hemoglobin variants that includes not only the common ones (like S and C) but also those that elude the usual techniques, such as Hb Lepore, would increase the impact of the proposed technique.

In the case of myeloproliferative neoplasms, there are interesting projects that would nicely complement the results presented here. First is to find the spectral differences between MPN and reactive non-MPN conditions as well as identifying their origin. This includes both SERS and DNA expression measurements. In addition, a broader population that includes myeloid leukemia patients, to test the capability of SERS to discern between MPN patients and leukemia patients, would expand the applicability of SERS spectroscopy.

An immediate step for MPN diagnosis using SERS would be to develop regression models that allow a *blind diagnosis* of samples. A first step towards this goal was developed in Chapter 4 with the use of ROC curves that led to the definition of thresholds that optimized both sensitivity and specificity. However, more robust models like partial least square regression or neural networks could be implemented. Such models would need to be tested in a *blind experiment*.

Another possibility for broadening the scope of the research presented here may be that of substrate optimization. Many techniques are currently available to synthesize

nanoparticles in more controlled ways than that of the Tollens reaction. For example, lithography techniques allow for manufacturing of patterned substrates. A highly reproducible substrate will contribute to better reproducibility of the measurement, and consequently it would increase the overall performance of the technique.

Finally, the broad background emission described in Chapter 4 has been interpreted as enhanced fluorescence from certain cellular components. However, similar background spectra are observed from other SERS measurements, even at the single molecule level where quenching of the fluorescence is expected. A better understanding of the origins of the SERS background from cells would be a an important step towards the development of new and better sensing application schemes for SERS and other surface-enhanced optical processes.

6.3 References

[1] K. Wang, "Advanced imaging in GI mucosal disease: Do you see what I see?" Gastrointest Endosc **73**, 204–205 (2011).

[2] L. Wong Kee Song, S. Banerjee, D. Desilets, D. Diehl, F. Farraye, V. Kaul, S. Kethu, R. Kwon, P. Mamula, M. Pedrosa, S. Rodriguez, and W. Tierney, "Autofluorescence imaging," Gastrointest Endosc **73**, 647–650 (2011).

[3] N. Thekkek and R. Richards-Kortum, "Optical imaging for cervical cancer detection: Solutions for a continuing global problem," Nat Rev Cancer **8**, 725–731 (2008).

[4] J. Hegyi, V. Hegyi, T. Ruzicka, P. Arenberger, and C. Berking, "New developments in fluorescence diagnostics," J Dtsch Dermatol Ges **9**, 368–373 (2011).

[5] V. Masilamani, K. Al-Zhrani, M. Al-Salhi, A. Al-Diab, and M. Al-Ageily, "Cancer diagnosis by autofluorescence of blood components," J Lumin **109**, 143–154 (2004).

[6] Y. Kang, M. Si, R. Liu, and S. Qiao, "Surface-enhanced Raman scattering (SERS) spectra of hemoglobin on nano silver film prepared by electrolysis method," J Raman Spectrosc **41**, 614–617 (2010).

[7] D. Qi and A. J. Berger, "Chemical concentration measurement in blood serum and urine samples using liquid-core optical fiber Raman spectroscopy," Appl Optics **46**, 1726–1734 (2007).

[8] G. Zhu, Y. Hu, J. Gao, and L. Zhong, "Highly sensitive detection of clenbuterol using competitive surface-enhanced Raman scattering immunoassay," Anal Chim Acta **697**, 61–66 (2011).

[9] V. Ranc, A. Stanova, J. Marak, V. Maier, J. Sevcík, and D. Kaniansky, "Preparative isotachophoresis with surface enhanced Raman scattering as a promising tool for clinical samples analysis," J Chromatogr A **1218**, 205–210 (2011).

[10] P. Kleinert, M. Schmid, K. Zurbriggen, O. Speer, M. Schmugge, B. Roschitzki, S. S. Durka, U. Leopold, T. Kuster, C. W. Heizmann, H. Frischknecht, and H. Troxler, "Mass spectrometry: A tool for enhanced detection of hemoglobin variants," Clin Chem **54**, 69–76 (2008).

[11] A. H. Ahmed, K. G. Nicholson, J. S. Nguyen-Van Tam, and J. C. G. Pearson, "Effectiveness of influenza vaccine in reducing hospital admissions during the 1989-90 epidemic," Epidemiol Infec **118**, 27–33 (1997).

[12] E. Hak, E. Buskens, G. A. Van Essen, D. H. De Bakker, D. E. Grobbee, M. A. J. B. Tacken, B. A. Van Hout, and T. J. M. Verheij, "Clinical effectiveness of influenza vaccination in persons younger than 65 years with high-risk medical conditions: The PRISMA study," Arch Inter Med **165**, 274–280 (2005).

[13] J. Duarte, M. T. Pacheco, A. B. Villaverde, R. Z. Machado, R. A. Zangaro, and L. Silveira Jr, "Near-infrared Raman spectroscopy to detect anti-Toxoplasma gondii antibody in blood sera of domestic cats: Quantitative analysis based on partial least-squares multivariate statistics," J Biomed Opt **15**, 047 002 (2010).

[14] W. Yan, H. Lin, L. Jinghua, Q. Dian, C. Anyu, J. Yi, G. Xun, L. Chunwei, H. Wen, and W. Hong, "Preliminary study on the quick detection of Acquired Immune Deficiency Syndrome by saliva analysis using surface enhanced Raman spectroscopic technique," Ann Int Conf IEEE EMBS **2009**, 885–887 (2009).

[15] J. Saade, M. T. T. Pacheco, M. R. Rodrigues, and L. Silveira Jr, "Identification of hepatitis C in human blood serum by near-infrared Raman spectroscopy," Spectroscopy **22**, 387–395 (2008).